

C. A. Gemignani, D. Grignani, E. Marullo
**Sperimentazioni di catalogazione digitale
di fonti documentarie e iconografiche**

Estratto da Reti Medievali Rivista, VIII - 2007

[<http://www.retimedievali.it>](http://www.retimedievali.it)



Firenze University Press

Sperimentazioni di catalogazione digitale di fonti documentarie e iconografiche

di Carlo Alberto Gemignani, Daniele Grignani, Eleana Marullo*

L'origine di questo contributo va ricercata nelle riflessioni svolte in vista della creazione di un archivio digitale funzionale alla ricerca storica, nell'ambito del Dipartimento di Storia Moderna e Contemporanea dell'Università di Genova e in particolare del Laboratorio di Archeologia e Storia Ambientale (LASA) e dell'Archivio Ligure di Scrittura Popolare (ALSP). I primi parziali risultati e le problematiche incontrate durante la realizzazione del progetto, tuttora in corso di svolgimento¹, saranno esposti nella seconda parte del presente saggio (par. 2). Si è ritenuto opportuno premettere, tuttavia, una riflessione di più ampio respiro sulle tendenze che si sono di recente sviluppate, nel sistema della conservazione e della catalogazione delle fonti (non solo iconografiche), in seguito alla diffusione dei *software* Open Source (par. 1).

1.1 Introduzione

La progressiva espansione del sapere collettivo, grazie all'attitudine a trasferire all'esterno di sé facoltà sempre più alte, secondo André Leroi-Gourhan ha imposto all'umanità lo sviluppo di mezzi più idonei alla custodia di grandi quantità di informazioni. E la definizione di un corpo di conoscenze organizzate secondo criteri di accessibilità, conservate e tramandate attraverso dispositivi di memorizzazione esterni, costituisce il presupposto per il passaggio dalla conoscenza individuale a quella collettiva praticato su vasta scala². Come è noto, il processo ha subito un'accelerazione a partire dal secolo XVIII, quando le possibilità offerte (da tempo) dalle nuove tecniche di

* Daniele Grignani è autore di tutto il paragrafo 1 e dei paragrafi 2.1, 2.2; Carlo Gemignani dei paragrafi 3.1, 3.2; Eleana Marullo dei paragrafi 3.3 e 3.4.

stampa crearono le condizioni per una sua maggiore distribuzione che, da patrimonio riservato, iniziò davvero a diffondersi. Ma è stato nel secolo XX che la quantità d'informazione ha raggiunto volumi tali da riempire sterminate biblioteche, mentre al contempo è cresciuta la necessità di organizzare i dati mediante strumenti supplementari in grado di garantire una più alta libertà di consultazione parcellare semplificata al massimo.

Dal momento che le conoscenze, per evolvere, devono trasformarsi in elementi documentari suscettibili di ricostruzioni appropriate e di arricchimenti indefiniti non volatili, nel 1945 Vannevar Bush immaginò il Memex (*memory expansion*), ovvero uno strumento col quale superare i limiti di strumenti di memorizzazione per così dire inerti: sfruttando la tecnologia dei microfilm, esso avrebbe consentito la memorizzazione di grossi archivi di informazioni e libri, mentre la presenza di strumenti di interazione per l'aggiunta di note e la creazione di collegamenti stabili tra documenti diversi, lo avrebbero reso capace di evolvere in modo indefinito³. Il dispositivo non vide mai la luce⁴, ma la comparsa dell'informatica, di nuove forme di telecomunicazione di massa e del *network-based computing* su larga scala hanno fatto compiere un decisivo passo avanti ai dispositivi di memorizzazione esterna, i cui contenuti ora possono prestarsi a collegamenti infiniti e a una maggiore libertà di consultazione.

Gli effetti di tali tecnologie sembrerebbero materializzarsi ben al di là della semplice divulgazione: anzi parrebbero agire nello stesso modo in cui l'incremento di giornali, riviste ed enciclopedie, "popolari" o meno, agirono nella Gran Bretagna del Settecento, quando la diffusione di «frammenti di conoscenza» finì con lo stimolare un interesse quasi generale verso la meccanica e la tecnica⁵. Ciò provocò l'innescare di quel processo di «solidarietà culturale e coesione sociale»⁶, causato dal possesso di un vocabolario tecnico e da interessi comuni, che permise la formazione dei cosiddetti "collegi invisibili": comunità di tecnici che collaboravano «attraverso una rete caratterizzata da un certo grado di connettività»⁷ scambiandosi informazioni ottenute per mezzo dei mezzi di comunicazione a loro disposizione.

Allo stesso modo, le nuove tecnologie sono alla base di un tipo di organizzazione informale⁸ in grado di mobilitare un grande numero di intelligenze che si caratterizzano per vocabolario, estrazione sociale e curiosità verso la tecnologia, e sono tra di loro in rapporto non tramite una rete fondata sulla prossimità fisica, bensì grazie a una infrastruttura tecnologica in grado di offrire un altissimo livello di connettività. E in questi ambienti, almeno all'inizio della loro storia, è possibile individuare una medesima tendenza ad affrontare problemi tecnici non più legati alla meccanica come nei secoli XVIII-XIX, bensì all'informatica, alla programmazione e alla creazione di *software*⁹.

1.2 L'Open Source

L'organizzazione informale cui sopra si accenna è nota oggi come Open

Source¹⁰. Fondata su un efficace modello di collaborazione e condivisione di informazioni, ha ben presto acquistato un ruolo di primo piano attraverso il coinvolgimento di migliaia di programmatori, mentre il *software* pubblicato trova impiego nell'industria e nelle università. I suoi successi, sempre più frequenti, dipendono da un ammasso di informazioni tecniche in parte codificate e in parte trasmesse in modo informale, attraverso la Rete, sotto forma di «questo funziona e questo no; questo si fa così». Molti dei prodotti più diffusi e noti si basano su un'idea iniziale vaga; mano a mano che sono state messe in pratica le nozioni tecniche necessarie (apprese all'università, sul luogo di lavoro e nella Rete) per costituirli, ci si è fatta un'idea di come operare fino alla costituzione di un embrione funzionante, dal quale si parte per capire come renderli più efficienti o per adattarli a un uso diverso.

Ma le nozioni tecniche rappresentano ancora una conoscenza estranea alla maggior parte degli utenti, anche se hanno assunto un significato nuovo dal momento in cui la cultura informatica è penetrata a ogni livello e i suoi specialisti non sono più confinati all'interno di "riserve" impenetrabili a causa di un linguaggio e di pratiche spesso incomprensibili. Allora la diffusione del *software* Open Source non appare più casuale: dopo anni di sporadico utilizzo di sistemi concepiti nella maggior parte dei casi per affrontare problemi di ordine eminentemente tecnico e pratico, o per aggirare gli alti costi e le imperfezioni del *software proprietario*, l'incontro tra competenze diverse ha prodotto un'evoluzione che ha coinvolto dapprima ambiti circoscritti e successivamente settori sempre più vasti del lavoro e del sapere.

In questo senso, è interessante rilevare come il modello collaborativo attuato dapprima all'interno delle facoltà universitarie si sia riversato nel Web, luogo nel quale ha superato i confini della tecnica in senso stretto dando vita a progetti come Wikipedia che, nonostante i limiti, rappresentano non solo un tentativo riuscito di gestione collettiva e informale dei contenuti, ma anche uno sforzo multidisciplinare capace di coinvolgere esperti di vari argomenti del sapere e tecnici in grado di rispondere alle loro esigenze¹¹. Ma ciò che rende l'Open Source interessante dal nostro punto di vista è il fatto che a differenza di quello *proprietario*, il *software* dal *codice sorgente aperto* nasce nelle università e nei centri di ricerca per rispondere a esigenze specifiche, e da lì si diffonde insieme alla pratica di collaborare, in questo caso per affrontare da sé, quando possibile, i problemi lasciati insoluti dall'impresa privata.

1.3 La storia e l'informatica

Gli sviluppi cruciali dell'informatica, infatti, hanno per lo più trascurato settori economicamente poco rilevanti, incapaci per loro natura di fornire un'adeguata prospettiva di profitto. E mentre l'uso di tecniche informatiche per la scrittura dei testi e la diffusione dei risultati della ricerca è ormai generale, il contributo fornito alla ricerca storica da *software* concepiti appositamente allo scopo è attualmente sporadico e occasionale. Questo accade

probabilmente perché il tipo di informazione trattata possiede attributi che influenzano profondamente la scelta degli strumenti *software* più adatti, anche se in passato l'introduzione di metodi sperimentali attuati per mezzo del calcolatore ha permesso il compimento di progetti considerati ormai "classici"¹², nei quali, come è noto, la tendenza a prendere in considerazione fonti omogenee ha avuto conseguenze soprattutto sulla quantità di dati trattati.

La ricerca storica, al contrario, parrebbe procedere attraverso un variegato insieme di informazioni di genere diverso, sospese entro un sistema nebuloso, fitto di relazioni difficili da ricostruire e ancor più da connettere, la cui catalogazione è profondamente influenzata dal rapporto che lo specialista stabilisce con le proprie fonti¹³. Perciò il modo in cui queste ultime verranno trattate per mezzo del computer dipende in larga misura dall'analisi preliminare del materiale raccolto e dalla conseguente organizzazione del dispositivo destinato ad accogliere gli attributi giudicati pertinenti. In realtà l'operazione è meno semplice di quanto possa apparire a prima vista poiché l'organizzazione della base di dati, la quale riflette in qualche modo i limiti di un'analisi preparatoria, è il risultato di un processo di astrazione che potrebbe essere considerato perfetto solo se l'intero corpo documentario fosse stato profondamente analizzato¹⁴.

La comprensione di questi meccanismi, e l'esatta percezione della complessità delle fonti, spinsero Manfred Thaller a elaborare un progetto dal nome di CLIO che, nonostante fosse costituito essenzialmente da un *database* (allo stesso modo di quelli che lo precedettero), ebbe implicazioni di una certa portata. Partendo dal presupposto che le procedure di digitalizzazione delle fonti rappresentavano (e rappresentano tuttora) un'attività particolarmente onerosa, il *software* da lui concepito, procedendo attraverso la separazione fisica tra le informazioni, il modello descrittivo e le regole necessarie alla loro elaborazione, avrebbe dovuto funzionare sviluppando la capacità di immagazzinare e di mettere in correlazione tra loro dati provenienti da fonti diverse. In questo modo sarebbe stata possibile l'immissione di qualunque tipo di attributi desunti dal materiale originale, ma senza la necessità di ricorrere ad approfondite analisi preliminari. Di conseguenza eventuali modifiche in corso d'opera, localizzandosi esclusivamente a livello di regole e modello, non avrebbero alterato in modo significativo una struttura concepita nel modo sopra descritto¹⁵. Così il patrimonio di informazioni non solo si sarebbe arricchito col tempo di un sistema di dati disponibile a tutti gli utenti, ma sarebbe stato pienamente suscettibile di ampliamenti, modifiche e manipolazioni successive.

Il progetto divenne operativo nel 1985 e, nonostante possedesse le caratteristiche descritte e fosse stato pensato per concentrare in sé differenti moduli operativi in grado di svolgere calcolo statistico, rappresentazione cartografica, conversione dei sistemi di misurazione o monetari in modo da «abbandonare un programma applicativo per passare a un altro» senza procedere alla sua completa rielaborazione¹⁶, non ebbe un grande successo.

Probabilmente un'interfaccia a caratteri che impone l'uso di comandi articolati e complessi¹⁷, nonché la conoscenza di alcuni principi base del *software* di non immediata comprensione, ne hanno compromesso la diffusione. Ma se da un lato CLIO ha rappresentato una tappa importante del processo di definizione di un *database* condiviso, capace di tenere conto di tutte le fonti disponibili, dall'altro la diffusione di Internet, la conseguente esplosione dell'Open Source, seguita da altre forme collaborative di incremento della conoscenza, costituiscono un altro tassello fondamentale per la concettualizzazione di un dispositivo *software* davvero utile al lavoro dello storico.

1.4 Depositi istituzionali

La diffusione del software Open Source in ambiti di norma ignorati dai prodotti commerciali e la sua successiva espansione al di fuori di questi, nonché l'“affermazione egemonica” di Internet e di tecnologie adatte a manipolare l'informazione in modo da renderla suscettibile di trasformazioni e arricchimenti indefiniti, hanno fatto sì che il modello collaborativo scavalcasse i confini entro i quali si era sviluppato; con la complicità di nuove tecnologie, porzioni sempre più vaste di utenti della Rete lo hanno adottato come orientamento generale. E se da un lato la comparsa dei cosiddetti programmi di *peer-to-peer*, nati per la diffusione di risorse spesso protette da *copyright*, ovvero filmati, musica, manuali in formato elettronico, ha cambiato le regole di fruizione di un certo tipo di media, dall'altro abbiamo assistito a una vera trasformazione degli strumenti di accesso alla conoscenza in direzione di una maggiore cooperazione per la sua creazione e divulgazione, le cui conseguenze più vistose sono forse la formazione di comunità reticolari impegnate in attività diverse e i tentativi più o meno volontari di aggirare i tradizionali canali della comunicazione e della divulgazione scientifica.

Nel frattempo si è giunti alla consapevolezza che buona parte dell'odierna produzione intellettuale, e più in generale porzioni rilevanti della nostra conoscenza, sono custodite sotto forma di documenti di tipo digitale caratterizzati da una grande varietà di formati. E com'è ovvio questa massa di dati ha iniziato a sollevare grossi problemi economici di conservazione, conseguenza del rapido invecchiamento delle piattaforme (di fatto ampiamente favorito e indotto dall'industria privata) al quale deve essere sommato il deterioramento dei supporti più diffusi. Per dare una risposta a queste esigenze, sono stati introdotti i cosiddetti *institutional repository*¹⁸: complessi *database*, profondamente influenzati dalla cultura della condivisione, pensati per la conservazione e divulgazione dei documenti in essi custoditi, quindi dotati di interfaccia web, di strumenti di *workflow* (vale a dire il processo di inserimento), di indicizzazione del testo, di controllo e modulazione degli accessi ai documenti e infine caratterizzati dalla presenza di un ambiente di catalogazione fondato sui *metadati*¹⁹.

Al momento la disponibilità è ampia²⁰; ma la maggior parte di essi,

nonostante le potenzialità delle tecnologie, non sono nulla più che ottimi strumenti di conservazione e divulgazione, mentre potrebbero operare sfruttando a fondo le possibilità offerte da soluzioni tecnologiche in grado di garantire un maggior livello di cooperazione tra gli utenti. Infatti banche dati così definite possiedono alcune peculiarità interessanti agli occhi dello storico proprio a causa della flessibilità basata su criteri di catalogazione degli attributi di un documento che potrebbero essere validi anche per questo tipo di ricerca. Tali sistemi hanno tuttavia trovato un maggior impiego in campo archivistico, mentre nel caso specifico, cioè la ricerca storica, è necessario separare i metodi del processo di ordinamento archivistico da quelli più propriamente storici che, imponendo il rispetto del dinamismo del “principio di provenienza”, influiscono pesantemente sulle tecniche di elaborazione dei sistemi informativi²¹.

Un esempio è costituito dai *server* E-Prints, ovvero «depositi di documenti digitali di ricerca, alimentati attraverso il processo noto come *self-archiving* o auto-archiviazione» all'interno dei quali trovano posto raccolte di documenti digitali prodotti durante le diverse fasi dell'attività di ricerca. Dall'embrione dell'idea, alla sua definizione, fino alla produzione del testo finale, ossia dal *pre-print* al *post-print*²², il modello si fonda su di un ciclo di produzione dell'informazione che oggi vede sempre più coinvolti i fruitori finali raccolti in comunità collaborative, integrate tra loro attraverso i nuovi mezzi di telecomunicazione di massa. Ma trattandosi più che altro di sistemi concepiti per la conservazione e la divulgazione, gli schemi di catalogazione nonché il funzionamento generale riflettono come vedremo, con rigida precisione, gli scopi per i quali sono stati progettati. Tuttavia, attraverso l'adozione e soprattutto l'adattamento dei versatili strumenti presi in considerazione, potrebbe essere possibile giungere alla definizione di un dispositivo efficiente e soprattutto utile al lavoro dello storico «senza pretendere che i ricercatori imparino i linguaggi della babele informatica [...]». Altrimenti rischiamo di far fare alle banche dati la fine dei vecchi inventari. I ricercatori le interrogheranno, le troveranno troppo complicate e poi torneranno dal consulente [...] a chiedere aiuto»²³ come poteva avvenire, per esempio, nel caso di CLIO.

1.5 I metadati e il software

L'Open Source non costituisce solo una fonte di *software* a buon mercato, bensì qualche cosa di più: un modello nel quale il lavoro si fonda su una collaborazione resa possibile dalle tecnologie disponibili. Il processo di per sé rappresenta un trasferimento di conoscenze individuali che diventano collettive nel momento in cui vengono incorporate nel *codice sorgente* diffuso attraverso vere e proprie “banche dati” accessibili via Web, capaci di mantenere inalterate alcune caratteristiche dei contenuti riconducibili al “principio di provenienza”. E dal momento che l'attività di collaborazione coinvolge talvolta centinaia di persone con idee e competenze diverse, le strategie di

registrazione delle fasi dell'opera assumono un ruolo cruciale all'interno del processo di produzione. In altre parole, deve essere sempre possibile ricostruire l'intero ciclo di vita del contenuto del *codice sorgente* attraverso un insieme di attributi custoditi dentro un *database* (più propriamente *repository*) che permetta la ricostruzione della storia e dell'uso e le relazioni degli elementi di cui è composto il *software*.

Le procedure potrebbero non essere dissimili da quanto imposto dalle necessità dello storico, il quale molto spesso si trova a dover intervenire su fonti tradotte in formato digitale caratterizzate da un insieme di qualità ben più complesse e di difficile identificazione, ma soggette a un dinamismo altrettanto rapido, soprattutto se sottoposte a continue manipolazioni causate dalla ricerca. Nel caso specifico, l'elemento chiave è costituito dai *metadati*, coi quali è possibile raccogliere e organizzare l'insieme delle caratteristiche di un documento: essi rappresentano una descrizione comprensibile alla macchina effettuata attraverso elementi standard, per la quale però è necessario costruire opportune procedure di elaborazione. Allo stato attuale il *set* di *metadati* più diffuso è Dublin Core, pensato per la descrizione di qualsiasi tipo di documento digitale attraverso un limitato numero di campi utilizzabili dall'autore, dal curatore o dall'editore.

Per mantenere un certo grado di flessibilità, lo standard si limita a definire il nome del campo e il relativo contenuto nel modo più generico possibile; per esempio l'elemento *contributor* rappresenta «an entity responsible for making contributions to the resource» che può essere l'autore, un illustratore, un'organizzazione oppure un servizio automatico²⁴. Ma in alcuni casi la presenza dell'elemento non è sufficiente a rendere efficace la descrizione: allo scopo di fornire maggiore dettaglio, sono stati introdotti i *qualificatori*. E mentre l'elemento *contributor* non ne possiede, l'elemento *coverage* può rappresentare l'estensione sia temporale sia spaziale della risorsa e per assumere il significato appropriato, deve essere combinato con il qualificatore *spatial* oppure *temporal* (che lo indicano rispettivamente come contenitore di una descrizione geografica o temporale) secondo una notazione che per comodità, potrebbe essere la seguente: *coverage.spatial* oppure *coverage.temporal*²⁵.

Nel caso di risorse particolari, Dublin Core può rivelarsi insufficiente e talvolta i gestori delle banche dati tendono a personalizzare lo schema secondo specifiche elaborate in proprio. Poiché lo scopo è l'interoperabilità tra le piattaforme, ottenuta attraverso la definizione di dati il cui valore e significato è noto e concordato, verrebbe da osservare che il risultato non può che ostacolare lo scambio o l'esportazione delle informazioni. In realtà l'inconveniente non è poi così grave, almeno laddove è possibile operare opportune traduzioni verso lo standard; ciò che a prima vista potrebbe sembrare una forzatura arbitraria, probabilmente rappresenta un punto di forza del Dublin Core: tenendo presente gli scopi per i quali è stato creato, mantenendo quindi una certa capacità di tradurlo nello schema corrente, il processo di adattamento ad applicazioni particolari può essere effettuato

senza eccessivo sforzo. Un esempio eccellente è costituito dall'adattamento operato durante la progettazione del Portale Italiano della Cultura On-line del Ministero per i Beni e le Attività Culturali (PICO). Come illustrato nel documento di sintesi prodotto dal Gruppo di Lavoro della Scuola Normale Superiore di Pisa²⁶, infatti, alcuni elementi hanno subito delle alterazioni allo scopo di adeguarli ai contenuti del sito, modifiche che in ogni caso non pregiudicano un successiva trasformazione in direzione dello standard. Infatti come è possibile vedere nella tabella 1.5.1, l'adattamento è avvenuto a livello del qualificatore; secondo le specifiche elaborate dalla «Dublin Core Metadata Initiative», il *software* di interpretazione deve essere in grado di ignorare i qualificatori sconosciuti riconducendo il contenuto all'elemento. In poche parole, le informazioni contenute nel campo *contributor.performer*, devono essere assegnate all'elemento *contributor* privo di qualificatore poiché il qualificatore *performer* non appartiene allo standard.

DUBLIN CORE ELEMENTS	ELEMENT REFINEMENTS
Contributor	Collaborator Editor Performer Producer Responsible Translator

Tabella 1.5.1: L'intervento di adattamento è stato realizzato a livello dei qualificatori. L'elemento contributor nello standard Dublin Core non possiede qualificatori, mentre in PICO sono adottati quelli visibili nella tabella (fonte: Documento di sintesi del progetto tecnico-scientifico per il Portale della Cultura Italiana del Ministero per i Beni e le Attività Culturali <http://www.otebac.it/siti/realizzare/PICO_2_1sintesi.pdf>, p. 8).

I *metadati* non possono sostituire le funzioni di un comune *database*: casomai quest'ultimo dovrà essere interamente organizzato attorno a essi. La tabella destinata ad accogliere i dati delle fonti non sarà dunque elaborata sulla base degli attributi di queste, bensì si caratterizzerà per la presenza di un campo generico destinato ad accogliere gli elementi e i qualificatori Dublin Core, collegato a una seconda tabella (attraverso una *chiave primaria*) contenente i dati veri propri. Ma se in questo modo l'utente non verrà più ostacolato dall'incompletezza dell'analisi preliminare e avrà piena libertà di alterare lo schema degli attributi aggiungendone altri o modificando quelli esistenti senza dover intervenire sulla struttura del *database* stesso, dall'altro è pur vero che sarà costretto a ricondurre i propri dati alle definizioni più o meno vincolanti stabilite dallo standard, anche se, come abbiamo visto, esiste la possibilità di estenderlo. Un esempio del modello di *database* descritto è riassunto nella figura 1.5.1.

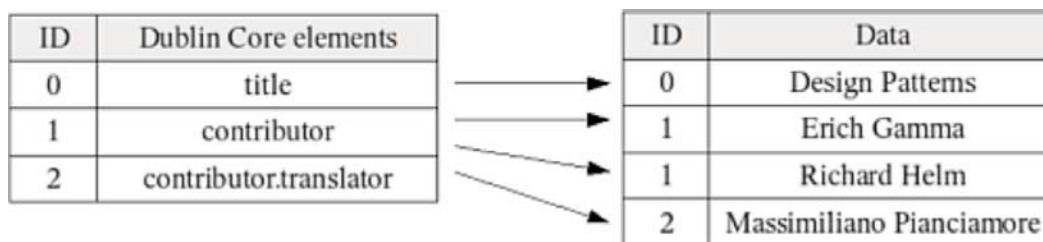


Figura 1.5.1: schema raffigurante le tabelle di un ipotetico *database* organizzato in base a un set di *metadati*.

Nonostante la natura puramente testuale del contenuto di entrambe le tabelle (a parte le *chiavi primarie*, nella figura rappresentate per mezzo dell'identificatore ID, memorizzate sotto forma di numeri interi, il cui scopo è quello di legare tra loro i dati delle differenti tabelle), la coppia elemento-qualificatore identifica il valore semantico del contenuto attraverso un formato adatto all'elaborazione automatica: l'elemento *title* rappresenta il titolo, *date* la data e così via. In alcuni casi, come quest'ultimo, lo standard prevede infatti una notazione particolare (anno-mese-giorno), mentre in altri, i termini della descrizione devono essere presi da vocabolari controllati, come per esempio l'elemento *type* (tabella 1.5.2), in grado di semplificare l'eventuale definizione di procedure o regole per il trattamento automatico dei dati.

DCMES Element	DCMI Type Vocabulary
Type	Collection Dataset Event Image Interactive Resource Moving Image Physical Object Service Software Sound Still Image Text

Tabella 1.5.2: L'elemento *type*, utilizzato per descrivere il tipo di risorsa, e l'elenco dei termini a esso associati definito dallo standard (fonte: DCMI Type Vocabulary, <<http://www.dublincore.org/documents/dcmi-type-vocabulary/>>).

La presenza di descrizioni soggettive ricondotte ad attributi oggettivi noti rappresenta un vantaggio anche nella fase di *information retrieval* non solo grazie all'uso di parole chiave opportunamente predisposte, ma soprattutto a causa degli elementi Dublin Core in grado di definire il campo semantico di ciò che descrivono. E nel momento in cui l'utente ha trovato ciò che cerca, se

abilitato, possiede l'opportunità di aggiungere nuovi valori agli attributi esistenti (un nuovo autore, per esempio), oppure nuovi attributi al documento visionato, magari nuove relazioni con altri presenti nello stesso archivio oppure trovati in Rete e così via. In altre parole tutte le informazioni rilevanti, catalogate per mezzo di definizioni oggettive, possono entrare a far parte del documento e quindi diventare patrimonio della comunità di studiosi, insieme a relazioni, qualità e descrizioni prodotte in tutti i casi in cui esso verrà studiato; anche se la coerenza di tutto questo dipende ovviamente dall'accuratezza dell'intervento umano, cioè dalle pratiche correnti all'interno dell'istituzione che si fa carico di gestire il sistema.

Sebbene la diffusione dell'uso dei *metadati* rappresenti un grosso passo in avanti, essi trovano impiego soprattutto all'interno di strumenti il cui funzionamento è riconducibile a quello di un archivio elettronico che sia privato dell'inerzia dei vecchi sistemi "analogici". La ragione di questo tipo di sviluppo deve probabilmente essere cercata nella tradizionale tendenza della biblioteconomia, caratterizzata sin dagli anni Sessanta dello scorso secolo dalla presenza di sistemi elettronici, a usare per la catalogazione dell'informazione bibliografica sistemi di *metadati* come MARC. Con l'aumento della produzione di informazioni digitali di tipo sempre più eterogeneo, la concettualizzazione delle "biblioteche digitali" e l'espansione della Rete e dei fenomeni che la caratterizzano, il passo verso lo sviluppo e l'adozione dei depositi istituzionali è stato breve. Ma nonostante la presenza di caratteristiche interessanti, l'utilizzo di questi sistemi nell'ambito della ricerca storica presenta alcune difficoltà che richiedono l'avvio di un processo di adattamento tecnologico in direzione di una maggiore flessibilità.

2.1 *DSpace: il deposito istituzionale attivo nel Dipartimento di Storia moderna e contemporanea*

Come annunciato in sede di premessa, nella seconda parte di questo saggio si darà conto dei progetti in corso presso il Dipartimento di Storia moderna e contemporanea dell'Università di Genova (DISMEC) in ordine a diverse tipologie documentarie (fonti scritte di età contemporanea, cartografia storica). Al di là del riferimento cronologico, i problemi epistemologici sottesi hanno una valenza generale, e possono essere utili anche per gli storici del medioevo.

Attualmente è attivo presso il DISMEC un dispositivo Open Source prodotto dal MIT di Boston e da Hewlett e Packard, noto col nome di DSpace Federation²⁷. Gli elementi che hanno orientato tale scelta vanno individuati nelle caratteristiche tecnologiche e architetture note e ben documentate, soprattutto diffuse in ambito accademico (dove sempre più spesso sono utilizzate come strumento didattico) nonché nell'indubbia fama dell'istituzione che si è fatta promotrice del progetto, nella quantità di installazioni operative²⁸ e in altre considerazioni di carattere più o meno tecnico²⁹. Ma al pari di altri dispositivi presi in analisi, DSpace è stato progettato per scopi

che potremmo definire archivistici: perciò la sua adozione ha richiesto un paziente lavoro di concettualizzazione, che ha avuto l'effetto di stimolare un dibattito, dal quale si è a sua volta innescato un processo di revisione tecnica del dispositivo in parte ancora in fase progettuale e in parte in via di realizzazione.

In primo luogo, la banca dati del DISMEC contiene documenti la cui organizzazione ha richiesto una ridefinizione nominale delle categorie archivistiche adottate dal *software* in oggetto, che in origine sembravano appunto riflettere gli scopi per i quali esso è stato prodotto: *community*, *sub-community*, *collection* e *information object*³⁰. Per ovviare agli inconvenienti legati a definizioni poco familiari alla comunità degli storici, queste sono state sostituite, come illustrato nella figura 2.1.1, con altre dotate di maggiore chiarezza: *archivio*, *fondo* e *carta*. E mentre le prime due rappresentano semplici contenitori di sotto-insiemi, la *carta* (o *collection*, secondo la qualificazione adottata da DSpace) costituisce l'elemento dentro il quale trovano

Struttura dell'archivio	Definizione DSpace	Definizione DISMEC
<ul style="list-style-type: none"> • Archivio Ligure Scrittura Popolare (ALSP) <ul style="list-style-type: none"> ◦ Fondi <ul style="list-style-type: none"> ▪ Famiglia Ferrari <ul style="list-style-type: none"> ▪ Autori diversi ▪ Ferrari Adele ▪ Ferrari Andrea ▪ Ferrari Angelo ▪ Ferrari Celeste ▪ Ferrari Costantino ▪ Ferrari Dorino ▪ Ferrari Elvira ▪ Ferrari Giovanni ▪ Ferrari Maria ▪ Ferrari Silvio ▪ Fondo Alloisio <ul style="list-style-type: none"> ▪ Alloisio Gerolamo 	<ul style="list-style-type: none"> Community Sub-Community Sub-Community Collection Sub-Community Collection 	<ul style="list-style-type: none"> Archivio ... Fondo Carta Fondo Carta

Figura 2.1.1: particolare dell'indice generale dell'archivio digitale del DISMEC. A sinistra in blu, l'elenco degli oggetti contenuti organizzato secondo un principio gerarchico; a destra, indicate dalle frecce, le relative denominazioni adottate da DSpace e dal DISMEC (fonte: DismecSpace, <<http://dismecspace.unige.it/dspace/community-list>>).

posto i documenti veri e propri. La pagina di consultazione (figura 2.1.2), possiede un insieme di funzioni in grado di agire sulla loro visualizzazione che, per mezzo dei pulsanti visibili in figura, può essere organizzata per titoli, autori, oppure per data di inserimento; è altresì disponibile uno strumento di ricerca, che opera per porzioni di testo all'interno dei *metadati* indicizzati dal sistema³¹.

DISMECspace > Archivio Ligure Scrittura Popolare (ALSP) > Fondi > Famiglia Ferrari >

Autori diversi

Home page della carta

In:

Cerca per

o sfoglia per

Iscriviti al servizio di notifica via e-mail di questa carta, per essere aggiornato giornalmente dei nuovi documenti pubblicati

Figura 2.1.2: particolare della pagina di consultazione di una *carta*. In figura, oltre a quelli per la visualizzazione dei documenti, sono raffigurati i pulsanti «Inserisci in questa carta» e «Iscriviti». Mentre il primo serve a dare inizio al processo di inserimento di un nuovo documento, il secondo permette l'iscrizione al servizio automatico di notifica della presenza di nuovi documenti nella *carta* (fonte: DismecSpace, <<http://dismecspace.unige.it/dspace/handle/123456789/8712>>).

Ma l'esplorazione dell'archivio può avvenire secondo modalità differenti: oltre a un elenco completo delle raccolte (figura 2.1.1), è possibile sfruttare un meccanismo di ricerca avanzata (figura 2.1.3) capace di agire sull'intero contenuto della banca dati, oppure dei singoli archivi, attraverso l'utilizzo dei campi di *metadati* indicizzati³² nonché di *keyword* opportunamente predisposte.

Cerca:

Cerca tipo:

Cerca per:

AND

AND

Figura 2.1.3: maschera di ricerca avanzata, nella quale il campo «Cerca» consente di selezionare l'archivio desiderato, oppure l'intero *database*, mentre i campi «Cerca tipo» permettono la selezione dei campi Dublin Core indicizzati (fonte: DismecSpace: <<http://dismecspace.unige.it/dspace/advanced-search>>).

La consultazione dei singoli documenti avviene attraverso una scheda descrittiva (figura 2.1.4) contenente la raccolta di attributi associati all'oggetto. Poiché si tratta di elementi Dublin Core, il *software* provvede a una loro traduzione in termini equivalenti, ma comprensibili, secondo norme predi-

sposte dall'amministratore del sistema (figura 2.1.6). Il documento descritto, come abbiamo più volte sottolineato, è da considerarsi un *information object*, ossia un potenziale aggregato di elementi diversi. Nella scheda illustrata nella figura 2.1.4, esso si compone una singola unità (qui definita come *item*), ma se fosse stato costituito da un maggior numero di documenti (per esempio più pagine dello stesso diario), come in effetti è nella realtà, nel riquadro grigio «Files in questo Item» posto nella parte inferiore dell'immagine, sarebbero apparsi tutti i *files* dell'aggregato.

[DISMECSpace](#) > [Archivio Ligure Scrittura Popolare \(ALSP\)](#) > [Fondi](#) > [Famiglia Ferrari](#) > [Autori diversi](#) >

Titolo: AutoriDiversi_028B
Titoli alternativi: EPISTOLARIO
Autori: FAMIGLIA FERRARI
Keywords: Emigrazione; vita quotidiana
Data di immissione: 24-Apr-2007
Editore: ALSP
Abstract: Lettere di saluto in cui prevale il bisogno di riconciliarsi a ciò che si è
Descrizione: La corrispondenza si svolge tra i figli, emigrati negli USA, e i genitori r
 51L
Appare nelle Carte: [Autori diversi](#)

File	Descrizione	Dimensione	Formato
AutoriDiversi_028B.jpg		652Kb	JPEG

[Guarda/Apri](#)

[Mostra tutti i metadati del documento](#)

Figura 2.1.4: scheda descrittiva che accompagna ogni documento contenuto nel sistema. In essa è visibile un insieme limitato di *metadati*, ossia quelli ritenuti più significativi per una rapida consultazione (l'intero corpo di attributi può essere visualizzato agendo sul pulsante «Mostra tutti i metadati del documento» collocato nella parte inferiore dell'immagine), nonché una raffigurazione dell'icona (nell'immagine: in basso a destra) attraverso la quale è possibile visualizzare il documento, in questo caso in formato JPEG (fonte: DismecSpace: <<http://dismecspace.unige.it/dspace/handle/123456789/9800>>).

Per quanto riguarda l'inserimento, l'azione procede lungo un sistema di *workflow* dipendente da modalità definite nelle singole *carte*, la cui sistemazione però è compito dell'amministratore dell'intero sistema. Il meccanismo, basato su un'interfaccia organizzata secondo un'idea di «scheda» (figura 2.1.5), comprende fasi che vanno dalla scelta della *carta*, alla compilazione di un ristretto insieme di attributi (titolo, autore e poco altro), quindi l'inserimento del documento vero e proprio (che può essere un file prodotto da Word oppure un video o altro), fino all'invio della richiesta di pubblicazione dell'oggetto all'amministratore della singola raccolta: soltanto dopo la revisione del documento inserito e degli attributi assegnati, esso potrà essere

pubblicato, ossia inserito nella banca dati e visibile agli utenti secondo le norme stabilite per la *carta* destinata ad accogliere il contributo (accesso pubblico o riservato).

The image shows a web form for submitting a document. At the top, a horizontal sequence of buttons represents the workflow: 'Descrivi' (highlighted in red), 'Descrivi' (bordered in red), 'Descrivi', 'Upload', 'Verifica', 'Licenza', and 'Completa'. Below this is a blue link: 'Submit: Descrivi il tuo Item'.

The form contains several sections with input fields and labels:

- Authors:** A section titled 'Enter the names of the authors of this item below.' It has two input fields. The first is labeled 'Cognome' with the example 'e.g. Smith'. The second is labeled 'Nome(i) + "Jr"' with the example 'e.g. Donald Jr'. There is an 'Aggiungi altro' button to the right.
- Title:** A single input field with the label 'Enter the main title of the item.'
- Series/Report No.:** A section titled 'Enter the series and number assigned to this item by your community.' It has two input fields. The first is labeled 'Series Name'. The second is labeled 'Report or Paper No.'. There is an 'Aggiungi altro' button to the right.
- Identifiers:** A section titled 'If the item has any identification numbers or codes associated with it, please enter the types and the actual numbers or codes below.' It has a dropdown menu labeled 'Identifiers' with 'ISSN' selected, followed by an input field. There is an 'Aggiungi altro' button to the right.
- Type:** A section titled 'Select the type(s) of content you are submitting. To select more than one value in the list, you may have to hold down the "CTRL" or "Shift" key.' It features a list box with the following options: Animation, Article, Book, Book chapter, Dataset, and Learning Object.

Figura 2.1.5: immagine raffigurante la scheda di inserimento del singolo documento. Nella parte superiore, è possibile vedere lo stato del *workflow*; in rosso l'azione corrente, mentre bordata dello stesso colore: l'azione compiuta; di seguito le altre ancora da effettuarsi. Ogni campo di testo, corrisponde a uno specifico elemento Dublin Core; «Authors» a *contributor.author*, «Title» a *title* ecc. La scheda ovviamente dipende dal contenuto del *registro dei metadati* e può essere liberamente modificata partendo da quest'ultimo.

Il sistema viene fornito con un insieme di *metadati*, mutuato dal Dublin Core, sul quale è possibile intervenire effettuando inserimenti oppure modifiche. Ma essendo tale insieme contenuto nel cosiddetto *registro Dublin Core* («Dublin Core Type Registry»)³³, i cambiamenti, permessi esclusivamente all'amministratore, avranno conseguenze sull'intero sistema; per esempio, se volessimo alterare l'attributo *contributor* (che in DSpace, al contrario di quanto accade nel Dublin Core standard, viene fornito con i qualificatori *advisor*, *author*, *editor*, *illustrator* e *other*) sostituendolo con un altro termine, finiremmo con l'alterare tutti i campi dello stesso tipo contenuti nell'archivio.

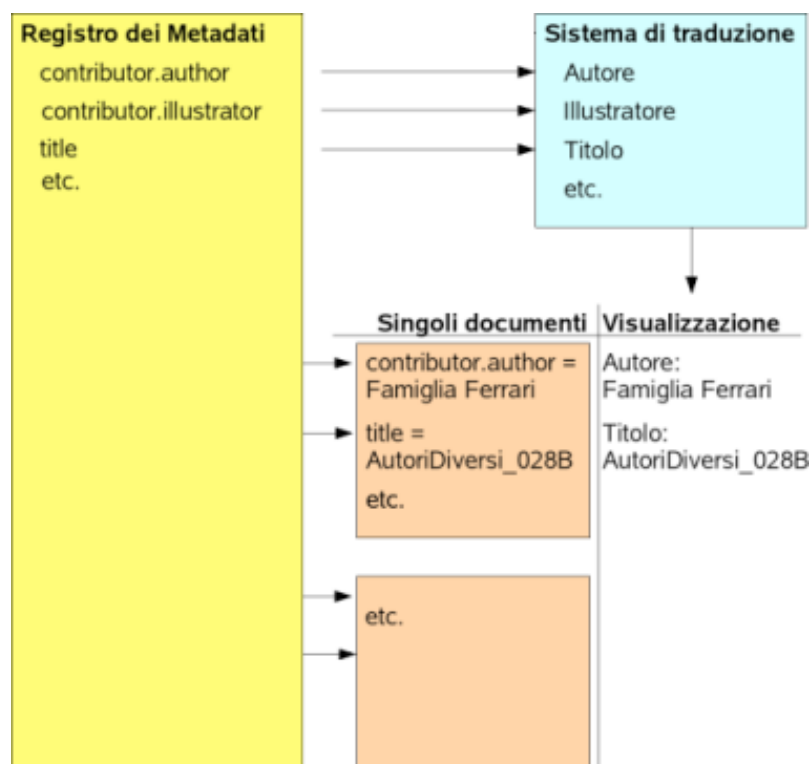


Figura 2.1.6: schema raffigurante il sistema di gestione dei *metadati*. In giallo, il registro dei *metadati*, contenente il *set* applicato all'intero sistema; in arancione la rappresentazione dei documenti, i quali prelevano i campi appropriati dal registro e li associano ai dati assegnati. In azzurro il sistema di "traduzione" del campo Dublin Core con la definizione umanamente comprensibile a esso assegnata.

Ciò non è altrettanto vero per i valori e i campi associati al documento. A tale livello infatti, è possibile effettuare variazioni senza che queste influiscano sul resto dei dati del sistema (figura 2.1.6). Perciò, come evidenziato nel paragrafo 1.5, è possibile intervenire sugli attributi del singolo documento aggiungendo elementi ricavati dal *registro dei metadati*, oppure modificando il contenuto degli attributi già presenti (figura 2.1.8).

Il processo di inserimento può avvenire sia nel modo precedentemente descritto, sia fornendo *metadati* e documenti, entrambi predisposti in un modo tale da essere compresi dal dispositivo. Come vedremo nel paragrafo 2.2, ciò offre l'opportunità di effettuare inserimenti in modalità *batch*, ossia di grandi quantità di informazioni alla volta. La particolare natura dei *metadati* li rende adatti a una rappresentazione in formato XML, come raffigurato nell'immagine 2.1.7. In essa purtroppo non è possibile distinguere gli elementi e i qualificatori elaborati in seguito alle esigenze identificate presso il DISMEC: la versione di DSpace in uso, la 1.3, non prevede l'utilizzo di un *prefisso di namespace* nelle espressioni XML.


```

<?xml version="1.0" encoding="utf-8" standalone="no"?>
<dublin_core>
<dcvalue element="contributor" qualifier="author">Rossi Anna</dcvalue>
<dcvalue element="subject" qualifier="none">Emigrazione</dcvalue>
<dcvalue element="title" qualifier="alternative">Epistolario Rossi</dcvalue>
<dcvalue element="coverage" qualifier="spatial">Bargagli (GE)</dcvalue>
<dcvalue element="coverage" qualifier="spatial">Stockton (California)</dcvalue>
<dcvalue element="date" qualifier="created">1914-1932</dcvalue>
<dcvalue element="format" qualifier="medium">M 058-058A</dcvalue>
<dcvalue element="publisher" qualifier="none">ALSP</dcvalue>
<dcvalue element="identifier" qualifier="sici">95</dcvalue>
<dcvalue element="identifier" qualifier="other">solo riproduzione</dcvalue>
<dcvalue element="description" qualifier="none">Sorella di Antonio, emigra in Argentina.</dcvalue>
<dcvalue element="source" qualifier="none">Copia digitale</dcvalue>
<dcvalue element="description" qualifier="dominanttypology">Epistolario</dcvalue>
<dcvalue element="description" qualifier="abstract">Anna Rossi dall'Argentina e Antonio, [...]</dcvalue>
<dcvalue element="language" qualifier="iso">it</dcvalue>
<dcvalue element="title" qualifier="none">001A</dcvalue>
<dcvalue element="type" qualifier="none">Image</dcvalue>

```

Figura 2.1.7: nella figura, la porzione (incompleta) di un file XML, contenente alcuni *metadati* e i rispettivi valori, visualizzato attraverso un programma simile al «Blocco note» (Notepad) di Windows. Esso è composto da *tag*, allo stesso modo del più noto HTML; in verde sono rappresentati i valori del campo Dublin Core (elemento più qualificatore), il cui nome è invece in rosso. In nero non grassetto: contenuto del campo, ossia il valore dell'attributo indicato dalla coppia elemento più qualificatore. Date le evidenti regolarità, un *computer* dotato del *software* necessario è in grado di popolare un *database* opportunamente predisposto partendo dalle istruzioni fornite.

Se ciò fosse possibile, a differenza di quanto è visibile in figura, la sintassi corretta per indicare un elemento appartenente allo standard, potrebbe essere qualche cosa di simile a: «<dc:value element="language" qualifier="iso">it</dc:value>»; mentre per indicare un elemento non standard, potrebbe essere: «<alsp:value element="birth" qualifier="place">Genova</alsp:value>». La presenza del *prefisso di namespace* («dc:» o «alsp:») renderebbe il codice decisamente più chiaro e leggibile (ed elaborabile).

type		Image	it	Rimuovi
birth	place	ROMFEGGIO (PC)	it	Rimuovi
identifier	bibliographiccitation	RELAZIONE DI SEMINARIO 1987-88	it	Rimuovi
description	dominanttypology	EPISTOLARIO	it	Rimuovi
education		ELEMENTARE	it	Rimuovi
<div> <div>amount</div> <div></div> <div></div> <div>Aggiungi</div> </div>				

Figura 2.1.8: porzione della pagina di modifica dei *metadati* assegnati al documento. Nella figura è possibile vedere solo una parte degli attributi realmente presenti; a sinistra sono elencati gli elementi Dublin Core (in questo caso sono ne sono visibili alcuni non appartenenti al *set* standard, ma elaborati presso il DISMEC: *birth.place* e *description.dominanttypology*) e il relativo qualificatore, quando esiste, mentre al centro è riportato il valore a essi assegnato. I pulsanti a destra permettono di rimuovere l'elemento Dublin Core e il suo valore. Nella parte inferiore dell'immagine, è visibile l'area riservata all'inserimento di un nuovo elemento desunto dal *registro dei metadati*. Ovviamente la pagina è raggiungibile soltanto dagli utenti abilitati (fonte: DismecSpace: <<http://dismecspace.unige.it>>).

Qui è bene ricordare che XML non è altro che un semplice documento di testo (l'unico formato caratterizzato da una certa stabilità nel tempo) strutturato secondo norme precise. Perciò tali documenti, non solo saranno umanamente leggibili attraverso un qualunque *editor* di testo (per esempio il noto "Blocco Note" di Windows) di qualunque sistema operativo, ma anche organizzati in modo da essere facilmente ri-assimilabili, attraverso l'intervento di un tecnico e l'elaborazione di procedure adeguate, da qualunque *database* attuale e futuro (figura 2.1.9).

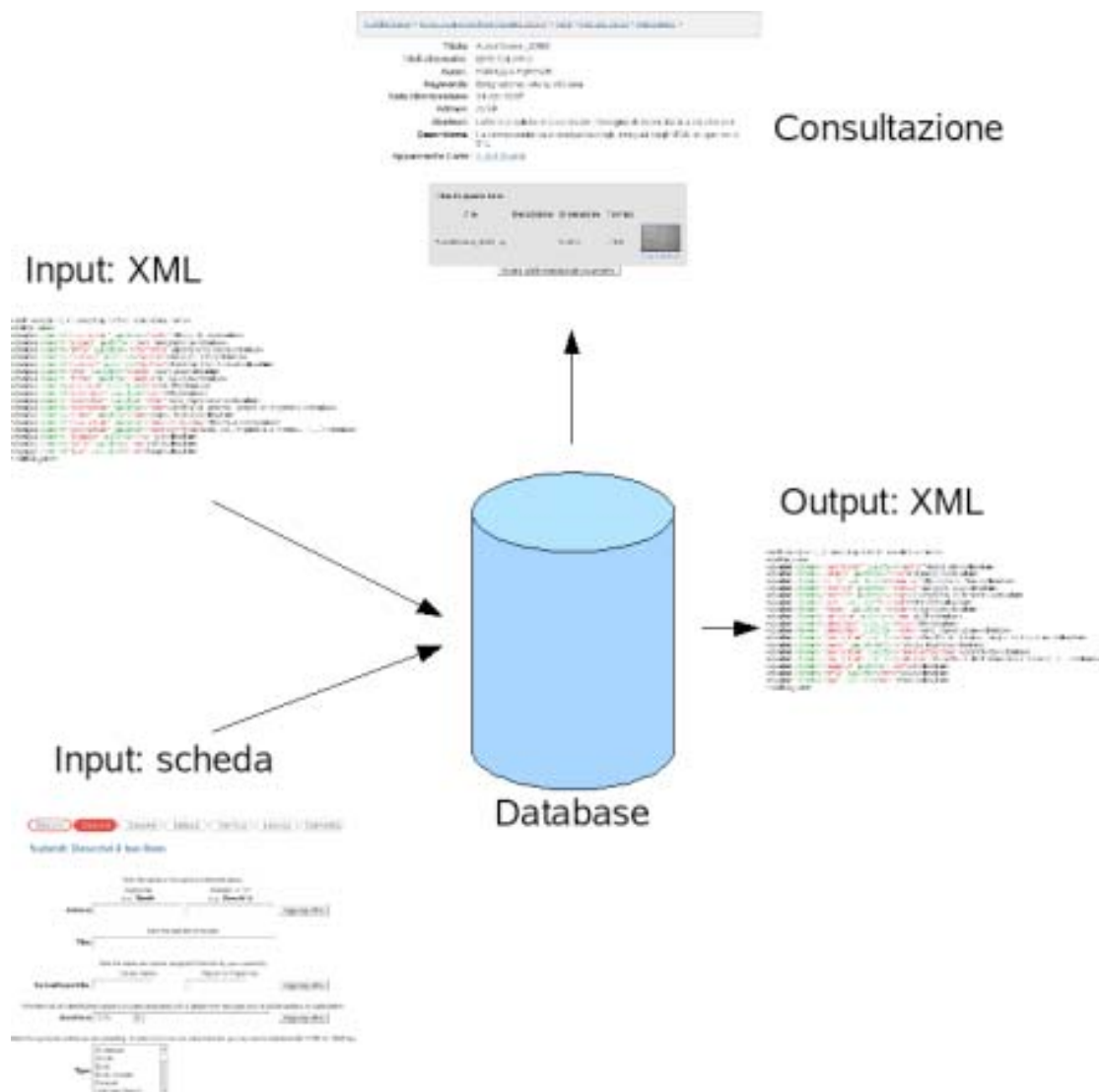


Figura 2.1.9: schema raffigurante i modelli di *input* e *output* dei dati in DSpace. Com'è possibile osservare, l'inserimento può avvenire per singolo documento attraverso la compilazione della scheda visibile in basso a sinistra, oppure direttamente in formato XML, purché compatibile con la struttura del *database*. In entrambi i casi, il dispositivo opera una conversione dei dati inserendo i valori forniti nelle tabelle appropriate, il cui contenuto viene opportunamente organizzato per la visione nel momento in cui l'utente opera una richiesta (in alto al centro in alto). Nel caso di esportazione dei dati, il dispositivo dispone di un efficiente meccanismo capace di estrarre i dati dal *database* e, in base alla loro organizzazione, di strutturarli in formato XML (a destra nell'immagine).

L'accesso ai contenuti, alle modifiche e più in generale alla consultazione della banca dati, può essere definito da un insieme di regole in grado di influire sul modello di fruizione di singoli documenti o di intere raccolte. Nell'uso reale è possibile immaginare la presenza di contenuti liberamente consultabili, mentre altri potranno essere resi accessibili soltanto ad alcune categorie di utenti. E tra questi ultimi, alcuni avranno la facoltà di modificare i *metadati* a livello di documento: come già sottolineato in precedenza, l'amministratore del sistema può assegnare al singolo o a gruppi di utenti, i pri-

vilegi necessari.

L'uso al momento è semplice: una volta stabilito il contenuto del *registro dei metadati*, l'utente avrà la facoltà di intervenire sugli attributi del singolo documento. Attraverso l'inserimento di nuovi elementi, descrizioni e relazioni, il ricercatore può intervenire in corso d'opera sull'insieme di attributi identificati e registrati in fase di analisi preliminare. Ma come sottolineato in precedenza, a causa degli scopi per i quali è stato progettato, ossia la catalogazione archivistica di documenti digitali, il dispositivo – utilizzato nella ricerca storica secondo un modello per così dire dinamico – ha rivelato alcuni limiti, qui brevemente riassunti:

1. il registro dei *metadati* è centralizzato: dunque il ricercatore che, nelle proprie fonti, individua attributi non contemplati nel *set* di elementi presente nel *registro dei metadati* per inserirli deve per così dire “inoltrare la richiesta” all'amministratore del sistema.
2. Le modifiche e gli interventi effettuati sul singolo documento si rifletteranno sull'intera raccolta, ossia saranno immediatamente disponibili a tutti gli utenti abilitati alla sua consultazione; questo potrebbe rivelarsi un problema dal momento in cui lo specialista opera sull'oggetto inserendo nel sistema attributi frutto di ipotesi suscettibili di modifiche.
3. Il dispositivo è basato su un sistema gerarchico organizzato in community (fondi), collection (carte) e information object (figura 2.1.1). Tuttavia è per così dire orientato al singolo documento, mentre ogni elemento (community, collection) dovrebbe poter essere considerato un oggetto informativo esso stesso, ovvero un aggregato composto da elementi diversi (per esempio: una community da diverse sub-community con le loro collection e relativi documenti), quindi descritto attraverso uno schema di metadati e trattato allo stesso modo dei singoli documenti.
4. Le gerarchie descritte nel punto 3 e illustrate nella figura 2.1.1 del presente paragrafo, una volta definite, sono irreversibili.
5. Le caratteristiche illustrate nel punto 3, possono altresì influire sulle modalità di ricerca, la quale avviene solo tra i documenti, mentre invece sarebbe utile poter disporre di uno strumento in grado di agire a ogni livello.
6. La ricerca complessa (figura 2.1.3) può avvenire utilizzando un interfaccia in grado di produrre query relativamente semplici, mentre al contrario in alcune occasioni sarebbe necessario disporre di uno strumento di interrogazione caratterizzato da un maggior livello di complessità.

7. Al momento DSpace permette ricerche attraverso un numero limitato di elementi del set di metadati registrato nel sistema (keyword, autore, titolo, abstract, lingua e argomento).
8. L'organizzazione gerarchica degli elementi che compongono gli archivi, è definita durante la loro fase di creazione, mentre i documenti troveranno posto nelle carte durante l'operazione di inserimento. L'operazione è rigidamente strutturata e, come sottolineato nel punto 4, irreversibile. In realtà sarebbe necessario disporre di uno strumento di aggregazione ben più flessibile, ovvero in grado di produrre gerarchie dinamiche basate sulle richieste dell'utente. L'operazione potrebbe essere possibile soltanto se ogni elemento dell'archivio fosse considerato un aggregato di documenti: un *information object* (punto 3).
9. Come già accennato nel paragrafo 1.5, le caratteristiche dei metadati potrebbero renderli adatti all'elaborazione automatica attraverso la creazione di moduli software opportunamente organizzati. Ma ciò imporrebbe la modifica del dispositivo in modo da renderlo adatto all'uso dei cosiddetti plug-in, ovvero di componenti esterni creati rispettando regole definite all'interno del software principale.

Forse il limite più vistoso è costituito dal principio secondo il quale l'*information object* si localizza esclusivamente livello del documento. Per Joaquim Carvalho «un'unica fonte dovrebbe essere trattata come un unico file»³⁴, e il concetto di *information object*, d'introduzione recente, sembrerebbe rispondere a questa esigenza. Ma in DSpace ciò si traduce in un'impostazione tecnica che lega l'aggregato a ciò che potremmo definire un'unica scheda (figura 2.1.4), mentre ogni pagina di un diario, di una lettera o di una memoria, costituisce un documento dotato di attributi e informazioni che devono essere in ogni caso registrati. Una buona alternativa, sarebbe quella di estendere il concetto sia alla *carta* che al *fondo*: solo in questo modo potrebbe essere possibile sfruttare a fondo le opportunità offerte da questo tipo di sistemi.

I *metadati*, e in particolare Dublin Core, rappresentano una base potenziale per la descrizione e il trattamento di riproduzioni digitali delle fonti storiche. Non solo essi forniscono l'interoperabilità tra le piattaforme *software*, ma la loro importanza è costituita dalla facilità con la quale possono essere utilizzati per la descrizione di un grande numero di risorse digitali: già da tempo sono operativi «linguaggi» basati sul paradigma dei *metadati* e utilizzati per la ricerca e lo scambio di informazioni, per la catalogazione, per la descrizione di risorse Web e molto altro ancora³⁵. Data la flessibilità della sintassi XML, essi potrebbero prestarsi all'elaborazione di un nuovo «linguaggio per la rappresentazione delle fonti», secondo la definizione data da Carvalho³⁶, dotato della stessa efficacia del *source oriented language* utilizzato da CLIO, ma rinnovato e calibrato ai nuovi mezzi di comunicazio-

ne, catalogazione e divulgazione. Sebbene molti settori siano giunti alla definizione di standard condivisi, in ambito storico la ricerca in tal senso, dopo il grande fermento del passato, sembra vivere un momento di torpore. Al contrario, proprio grazie alle esperienze pregresse potrebbe essere davvero possibile giungere alla definizione di uno standard di descrizione delle fonti digitalizzate, oppure all'adozione di un sistema già rodato e opportunamente adattato, a patto però di elaborare strumenti *software* appropriati alle esigenze degli storici.

Quanto all'attività operativa svolta presso il DISMEC, essa ha permesso di identificare alcune aree di intervento sulle quali è possibile operare allo scopo di migliorare le prestazioni di un strumento che, al momento, funziona più che altro come un ottimo dispositivo di catalogazione, conservazione e divulgazione. Gli adattamenti tecnici elencati sopra, concepiti per adeguare il *software* alle necessità della ricerca storica effettuata presso il DISMEC, sono attualmente in fase di progettazione avanzata. Ma la loro attuazione, essendo fondata su basi empiriche, non pretende di esaurire i problemi connessi all'uso di tali sistemi nell'ambito della ricerca. Al contrario, tali modifiche rappresentano un tentativo di definizione di un dispositivo efficiente, il cui compimento però dipende in larga misura sia da un dibattito dal quale possano scaturire le linee guida di una sua futura evoluzione, sia da risorse, umane ed economiche, difficilmente reperibili al di fuori della ricerca storica stessa.

2.2 L'esperienza dell'Archivio Ligure di Scrittura Popolare (ALSP)

Presso il Dipartimento di Storia Moderna e Contemporanea della Facoltà di Lettere e filosofia dell'Università di Genova, è da tempo attivo l'Archivio Ligure della Scrittura Popolare (ALSP), fondato e diretto da Antonio Gibelli allo scopo di «promuovere il recupero, la conservazione e lo studio – sotto diversi profili disciplinari – dei documenti di “scrittura popolare” e inoltre di sviluppare la ricerca in merito alla diffusione delle pratiche sociali di scrittura tra Ottocento e Novecento»³⁷. L'attività di ricerca svolta si fonda sulla raccolta di materiale documentario (diari, memorie, epistolari ecc.), rintracciato presso archivi privati (dunque acquisito), oppure riprodotto e custodito (in forma di microfilm, copia fotografica e digitale).

Il patrimonio complessivo di documenti è composto da oltre trentamila carte catalogate e gestite per mezzo di una vecchia, ma ancora efficiente banca dati. Ma l'intensa attività di studio e la continua manipolazione dei preziosi documenti, la deperibilità delle copie fotografiche e dei microfilm, unite alla necessità di registrare le scoperte effettuate a beneficio della comunità di studiosi e studenti raccolta attorno al centro, nonché le relazioni internazionali istituite con altre realtà assimilabili, hanno da tempo creato le condizioni per l'avvio di un più vasto e ambizioso programma di digitalizzazione dei documenti stessi. In seguito all'ampia disponibilità di materiale digitale provocata dall'iniziativa, si è resa necessaria la sostituzione del siste-

ma di archiviazione con un altro caratterizzato dalla possibilità di gestire direttamente le riproduzioni. E data la scarsità di letteratura tecnica in materia («questo funziona e questo no; questo si fa così»), il processo di migrazione si è articolato in fasi non del tutto distinte, ma qui schematizzate per comodità:

- 1) analisi del modello archivistico in uso (allo scopo di limitare, per quanto possibile, l'impatto sulle pratiche consolidate all'interno del centro);
- 2) identificazione delle necessità operative in relazione al budget e al tipo di utenza (abilità tecniche degli utenti), e conseguente ricerca del software più adatto;
- 3) analisi del tipo di documenti da catalogare, organizzazione del set di metadati (modifica e introduzione di nuovi elementi nell'ambito del Dublin Core, come descritto 1.5);
- 4) allestimento dell'archivio digitale, recupero e travaso dei dati dal vecchio al nuovo sistema, inizio dell'attività di inserimento dei documenti;
- 5) revisione dei dati provenienti dal vecchio sistema.

Durante le prove effettuate nel corso della fase descritta al punto 1, si è reso evidente che, a causa di processi legati alla presenza di *software* proprietario, il “travaso” dei dati da un sistema all'altro avrebbe comportato alcune difficoltà nella conservazione della loro integrità (come in effetti è stato; il risultato è visibile nella figura 2.1.4). D'altronde il mantenimento del vecchio dispositivo avrebbe compromesso l'attuazione del programma e, sul lungo periodo, la conservazione delle informazioni raccolte. Quindi si è scelto di operare con un sistema completamente Open Source che attualmente comprende il sistema operativo, i dispositivi di gestione del meccanismo di consultazione via Web dell'archivio, nonché la banca dati vera e propria. Il vantaggio è palese: a parte i costi (punto 2), data la loro particolare posizione di mercato questi *software* tendono a utilizzare procedure standard; e nel caso specifico il sistema di catalogazione è fondato sul *set* Dublin Core che, come è noto, viene riconosciuto dalla maggior parte degli strumenti di catalogazione più diffusi e al contempo è facilmente esportabile nel formato XML (del quale parleremo in seguito).

L'analisi delle caratteristiche dei documenti (punto 3) ha rivelato l'assenza di dati omogenei e la preponderanza di caratteristiche fortemente variabili, ossia totalmente dipendenti dalla natura dell'oggetto e dalle circostanze di redazione. Sicché, a parte un piccolo insieme di attributi ricorrente, esistono un grande numero di informazioni variegata, alcune delle quali compaiono solo occasionalmente, mentre altre possono riproporsi per un maggior numero di volte. Tali condizioni possono ostacolare la definizione di una “scheda” descrittiva sufficientemente esaustiva e soprattutto valida per ogni elemento, a meno di sacrificare alcuni dati a beneficio di altri ricorrenti.

Per Oscar Itzcovitch l'operazione è inevitabile: infatti lo storico, analizzando una fonte, opera una selezione preferendo tra le informazioni disponibili quelle che egli ritiene più significative³⁸. Ma se da un lato quelle raccolte e inserite in un ipotetico *database* rappresentano una fonte del tutto nuova, ossia costruita dallo storico stesso³⁹, è pur vero che lo schema utilizzato avrà una profonda influenza sull'organizzazione del dispositivo. Di conseguenza ogni variazione nella regolarità delle fonti non potrà che essere sacrificata a causa della semplificazione avvenuta all'interno della struttura del *database*. Al contrario, un sistema nel quale è possibile aggiungere liberamente commenti, annotazioni e attributi ricondotti a definizioni oggettive dovrebbe potersi collocare a monte del processo di selezione descritto: fonti specifiche, in seguito a nuove ricerche, potrebbero introdurre nuove entità e nuovi attributi allo schema di base senza che questo debba essere rivisto o manipolato, ma solo adattato alle nuove condizioni. In questo modo lo specialista possiede l'opportunità di selezionare dall'insieme informazioni conservate, le caratteristiche ritenute soggettivamente più significative⁴⁰.

Elemento Dublin Core	Qualificatore Standard	Qualificatore DSpace	Qualificatore ALSP	Qualificatore LASA
description	TableOfContents	provenance	OtherTypology	Ecological
	abstract	sponsorship	production	geographical
		statementOfResponsability	dominatTypology	historical
		tableOfContents		
		uri (uniform resource identifier)		

Tabella 2.2.1: un esempio di elementi appartenenti al *set* di *metadati* registrato nel DSpace installato presso il DISMEC. Come si può constatare, per l'elemento *description* lo standard Dublin Core prescrive appena due qualificatori.

Nel vecchio archivio informatico questi elementi erano presenti campi privi di un'esatta corrispondenza nello standard Dublin Core: perciò si è resa necessaria un'operazione di traduzione che ha imposto l'elaborazione di nuovi elementi da integrare a quelli esistenti (tabella 2.2.1). Sicché il trasferimento dei dati, come già accennato, si è rivelato un processo particolarmente gravoso a causa delle difficoltà legate all'incompatibilità tra i due sistemi. La perdita di informazioni è stata piuttosto consistente soprattutto nelle parti caratterizzate da una maggior quantità di testo; d'altra parte la revisione delle informazioni, dei refusi e in generale dell'accuratezza del contenuto del *database* avrebbero in ogni caso imposto un'attività di correzione generale.

La struttura del vecchio archivio digitale era organizzata secondo uno schema abbastanza semplice: tutto il materiale raccolto, suddiviso per fami-

glie (*fondi*) e al loro interno per scriventi (*carte*), era catalogato in base a ciò che di fatto costituiva l'unità minima in esso registrata: appunto la *carta*; lo schema è rappresentato nella figura 2.2.1. Ma come abbiamo visto, le nuove procedure hanno reso quest'ultima "soltanto" un contenitore al pari del *fondo*, mentre il singolo documento, ossia le pagine di un diario, di una lettera o di un ricettario, hanno acquisito un ruolo che potremmo definire centrale: in questo modo la scheda è stata condotta alla nuova unità minima: il singolo documento. Nella figura 2.2.2 è chiaramente visibile il modo in cui oggi si articola l'archivio digitale (rispecchiando fedelmente l'archivio per così dire analogico): il *fondo* è associato alla famiglia, mentre la *carta* rappresenta il contenitore dei singoli autori, i quali a loro volta rientrano in categorie definite a priori e aggiornate durante lo studio delle nuove acquisizioni (indicate dall'attributo *subject*): «Vita quotidiana», «Emigrazione», «Grande guerra» e via dicendo.

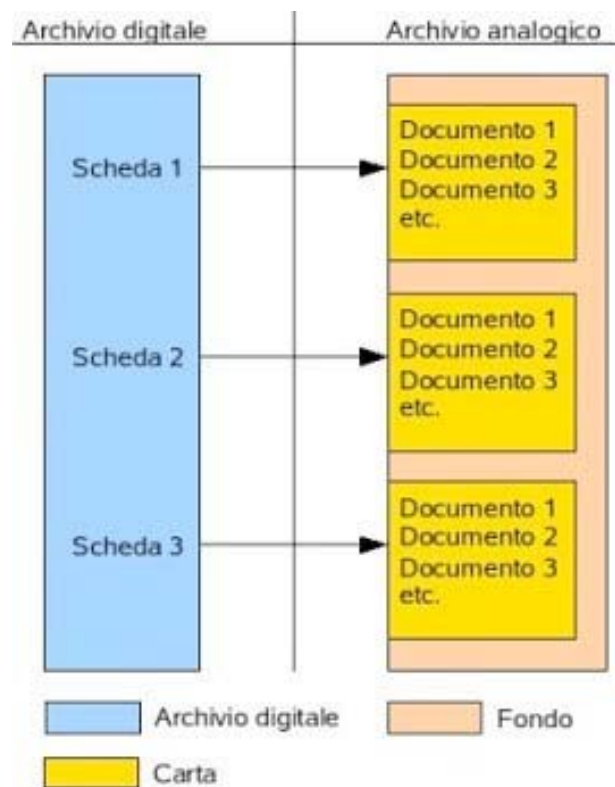


Figura 2.2.1: schema raffigurante il principio di catalogazione utilizzato nel vecchio archivio digitale. A sinistra è rappresentato il *database*, il quale appunto conteneva "schede" caratterizzate da campi descrittivi e la collocazione delle singole *carte*.

A causa delle differenze, la migrazione si è articolata in tre fasi distinte. Poiché da tempo è in atto una conversione in digitale del materiale raccolto, buona parte delle informazioni contenute nel vecchio archivio informatico non solo si riferivano a oggetti per così dire analogici, ma anche ad altri di natura appunto digitale. Quindi si è provveduto, in primo luogo, all'esportazione dei dati testuali delle singole schede in un formato che potremmo

definire di *interscambio*, ossia CSV (*Comma Separated Values*: valori separati da virgola) dal quale è stato poi ricavato un formato XML strutturato secondo il modello imposto dal nuovo sistema. Fatto ciò, e identificate le *carte* già digitalizzate, i *files* XML relativi sono stati per così dire “adattati” ai singoli documenti attraverso un processo eseguito da un *software* appositamente concepito per questo lavoro (seconda fase). Una volta ottenuta la normalizzazione del materiale, ossia la creazione automatica della scheda relativa al singolo documento e l’aggregazione in *carte* e *fondi* secondo un’organizzazione stabilita a priori, per facilitare il processo automatico di inserimento, si è proceduto infine alla creazione di un altro *software* in grado di importare tutto nel nuovo archivio⁴¹. Lo schema raffigurante l’attuale condizione è descritto nella figura 2.2.2.

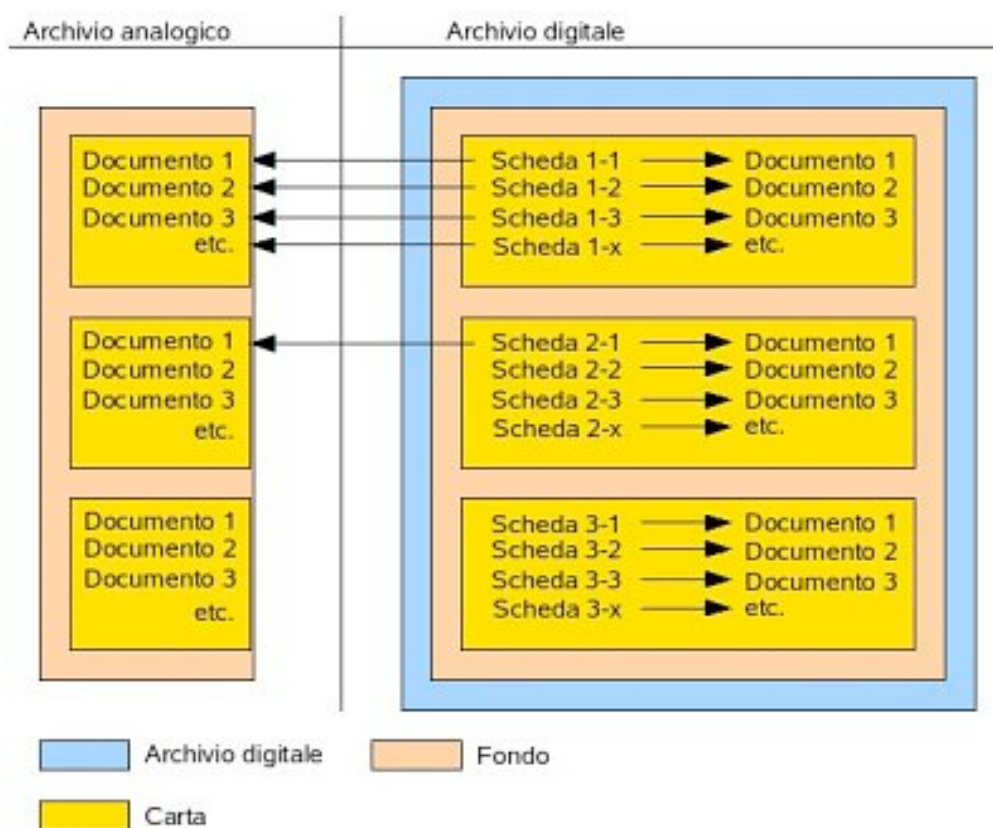


Figura 2.2.2: l’attuale condizione del nuovo archivio digitale. A sinistra è rappresentata la componente “analogica”, ossia l’archivio reale contenente i documenti cartacei, mentre a destra è raffigurato il dispositivo organizzato secondo lo schema descritto nel presente paragrafo. Come è possibile notare, ogni scheda contenuta in DSpace non solo contiene gli attributi dei documenti digitali, ma mantiene un riferimento ai documenti reali.

Attualmente in DSpace sono contenuti circa diecimila documenti, mentre il processo di inserimento è ancora in atto e contemporaneamente è iniziato il lavoro di revisione dei contenuti testuali. Anche quest’ultimo passaggio, come i precedenti, viene attuato attraverso procedure facilitate da un *software* concepito per operare con l’enorme massa di documenti

conservati dall'ALSP. Infatti ognuno di essi è contenuto all'interno di una *carta*, quindi è costituito da un insieme di elementi ricorrenti: descrizione dell'unità di appartenenza, argomenti trattati, autore e attributi relativi, collocazione archivistica, soggetto e altri ancora. Perciò la revisione del testo può avvenire su ciò che potremmo definire la "matrice" delle schede, ovvero gli elementi ricavati dalla rappresentazione XML della scheda iniziale ottenuta dal vecchio archivio ed estesa ai singoli documenti. Come abbiamo visto nel paragrafo 2.1 (figura 2.1.9), il sistema è dotato di un efficiente meccanismo di esportazione dei contenuti basato su XML. Una volta ottenuti tutti i *files* necessari, ossia tutti quelli contenuti in una *carta*, attraverso il *software* di correzione descritto è possibile intervenire contemporaneamente sui valori di tutti i documenti; sempre attraverso lo stesso *software*, i dati modificati possono essere nuovamente introdotti nel sistema sfruttando alcuni degli automatismi disponibili⁴².

Una delle conseguenze più importanti è costituita dall'opportunità di stabilire delle norme di conservazione, delle informazioni digitali, completamente indipendenti dalla piattaforma di custodia. Infatti il meccanismo di esportazione opera organizzando fisicamente il materiale digitale secondo lo schema più volte descritto: fondo/carte/documenti. L'equivalente all'interno del *filesystem* è costituito da una struttura fondata su *directory* e *subdirectory*, dove ogni documento è accompagnato dal *file* XML contenente la descrizione. Per cui automatizzando il processo di esportazione del contenuto del *database* su *hard disk*, è possibile ottenere una struttura come quella visibile nella figura 2.2.3, facilmente registrabile su CD oppure su DVD⁴³. E siccome non solo ogni elemento contiene tutti i dati necessari, ma l'intera struttura conserva l'organizzazione originaria (*fondi/carte/documenti*), è possibile assimilarla sia in un nuovo DSpace, sia in un altro dispositivo dello stesso genere opportunamente predisposto⁴⁴.

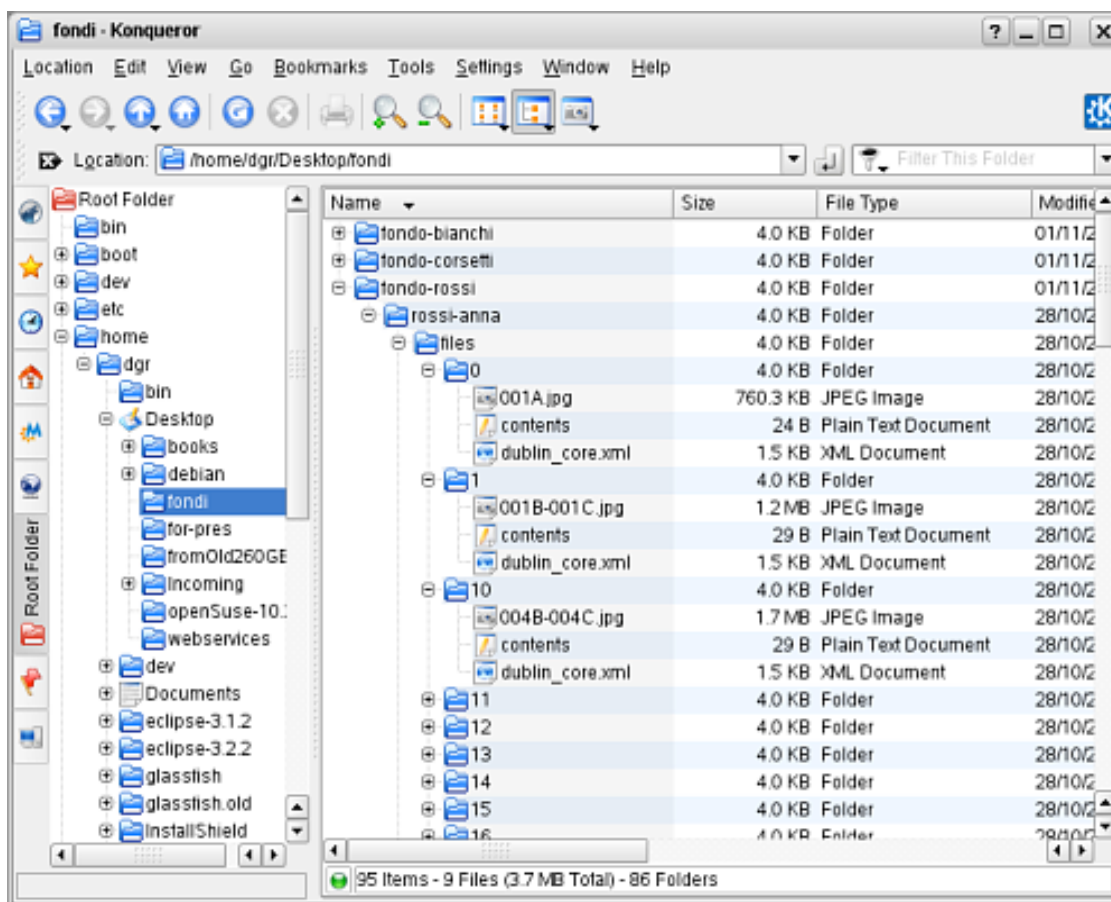


Figura 2.2.3: immagine raffigurante il risultato ottenuto esportando il contenuto del database. Ogni fondo contiene le proprie carte, mentre ogni carta contiene delle directory numerate nelle quali sono custoditi tre file: il documento vero e proprio (nell'immagine: file di tipo JPEG), un file di testo che indica qual'è il documento principale («contents»), infine il file XML contenente gli attributi del documento principale («dublin_core.xml»).

3.1 L'attività del Laboratorio di Archeologia e Storia Ambientale (LASA)

È in una prospettiva storico-analitica di approccio ai sistemi ambientali e territoriali⁴⁵ che, da oltre un decennio, opera il Laboratorio di Archeologia e Storia Ambientale (LASA, DISMEC-DIPTERIS Università di Genova). Negli ultimi anni le attività di ricerca⁴⁶, incentrate principalmente su una “rete di siti” localizzati nell'Appennino ligure orientale, si sono concentrate sulle problematiche legate all'individuazione e gestione del patrimonio storico-ambientale europeo.

La produzione e la valutazione di materiale archivistico riferito sostanzialmente a due tipologie di fonti: “di terreno”⁴⁷ e “documentario/testuali” (testi, immagini, registrazioni audio e video, ecc.), con la creazione di vere e proprie “monografie di sito”, ha sollevato i problemi di gestione di una consistente massa di dati, raccolti dai ricercatori secondo le specifiche competenze e nello svolgimento di specifiche ricerche, riversata su eterogenei formati digitali.

Dalle riflessioni sorte in merito alle necessità di “messa in relazione”,

“incrocio” e gestione dei dati è nata la proposta di sperimentazione di un *archivio digitale* di laboratorio, progettato per accogliere non solo materiale “storico” in senso stretto (riversato in copia su supporto digitale) ma anche i diversi risultati delle analisi compiute in base a esso, con la conseguente creazione di nuove “fonti”.

Il gruppo di lavoro costituitosi si è mosso a partire dall’analisi dei *software* di gestione di archivi attualmente più diffusi giungendo a identificare nei cosiddetti *institutional repository* un efficace dispositivo di organizzazione dei dati e di razionalizzazione del materiale in base alle priorità individuate: 1) conservazione nel tempo dei dati; 2) condivisione (interoperabilità, controllo) all’interno delle *équipes* di ricerca autorizzate all’accesso; 3) aggiornamento dei dati, per garantire la loro ri-discussione in base alle problematiche emergenti, creando così nuove occasioni di ricerca.

Il “contenitore”, che ha assunto il nome di DismecSpace⁴⁸, è stato messo a punto dopo prove e discussioni incentrate su due specifiche fonti, fotografia storica e cartografia, sulle quali si è concentrata una parte dell’attività di ricerca del LASA negli ultimi quattro anni.

3.1.1 La fotografia storica

Per quanto riguarda la fotografia storica l’importanza del documento, sempre legato a un dato spaziale e temporale preciso, è in relazione non solo alla sua antichità e alla possibile datazione, ma soprattutto alla possibilità di localizzazione del punto di ripresa (*spot*)⁴⁹. Questa operazione consente la verifica delle informazioni contenute nel documento attraverso il confronto con l’ambiente attuale (fotografie di confronto), con altro materiale originale di varia natura (per esempio la cartografia) e con i risultati di analisi specialistiche (archeologiche, geo-botaniche, geografiche, ecc.).

La prospettiva di attivare l’archiviazione su DismecSpace non solo del materiale storico ma anche, per esempio, delle riprese di confronto, ha fatto propendere per la sperimentazione di un *database* basato sull’integrazione di standard descrittivi riconosciuti (Dublin Core) quindi sull’elaborazione di una “scheda” descrittiva adatta alla fotografia che, senza trascurare l’identificazione delle relazioni archivistiche, se esistenti, tra l’oggetto fotografico “analogico” e la copia digitale, consentisse di valorizzare campi dedicati a specifiche esigenze di rilevamento.

L’esemplificazione è stata condotta sul sito del *Lago della Nava*, area posta a 1180 metri circa lungo lo spartiacque tra la val Trebbia e la val d’Ave-to (Appennino ligure orientale, comune di Fontanigorda, in provincia di Genova); inserito nel più vasto SIC (Sito di Interesse Comunitario) *Lago Marcotto – Roccabruna – Gifarco – Lago della Nave* (IT1331012 – Rete Natura 2000). La scelta è maturata in conseguenza dell’abbondanza di documentazione raccolta negli anni nella prospettiva della ricostruzione di un particolare paesaggio culturale: un alpeggio montano appenninico⁵⁰.

Alcune fotografie storiche relative al Lago della Nava, reperite presso

collezioni private e archivi istituzionali, sono state digitalizzate, schedate e inserite all'interno di DismecSpace⁵¹. Sulla base delle stesse immagini sono state effettuate alcune riprese digitali di confronto, anch'esse archiviate nel *repository*, che hanno completato la documentazione raccolta in seguito ai *fieldworks* effettuati tra la primavera 2006 e l'estate 2007.

Il tracciato adottato nella schedatura ha consentito di creare un tramite diretto tra le fonti storiche e le riprese attuali lasciando la possibilità di creare ulteriori e complesse relazioni con tutta la serie di dati raccolti.

Alcuni campi introdotti nella scheda, che accolgono informazioni utili alla ricerca storico-geografica ed ecologico-storica, sono stati tradotti in una forma compatibile con lo standard Dublin Core (tabella 3.2.1).

Tipi di dati	Campo	Descrizione	Dublin Core
Riferimenti minimi necessari per lo standard Dublin Core Metadata	Titolo	Nome attribuito alla risorsa (documento) digitale	<code><element>title</element></code> <i><code><!unqualified></code></i>
	Autore o creatore	Entità responsabile principale della produzione della risorsa	<code><element>contributor</element></code> <code><qualifier>author</qualifier></code>
	Soggetto e parole chiave	Descrizione del soggetto della risorsa ed eventuali parole chiave	<code><element>subject</element></code> <i><code><!unqualified></code></i>
	Descrizione	Descrizione testuale del contenuto della risorsa, incluso eventuale abstract	<code><element>description</element></code> <i><code><!unqualified></code></i>
	Editore	Entità responsabile della "diffusione" della risorsa	<code><element>contributor</element></code> <code><qualifier>editor</qualifier></code>
	Altro responsabile	Persona o ente che ha contribuito intellettualmente alla produzione della risorsa	<code><element>contributor</element></code> <code><qualifier>other</qualifier></code>

Data	Data in cui la risorsa è stata resa disponibile nella forma AA-MM-GG	<element>date</element> <qualifier>created</qualifier>
Tipo di risorsa	Categoria cui appartiene la risorsa	<element>type</element> <unqualified>
Formato	Formato dei dati, usato per identificare software e hardware necessari per utilizzare la risorsa	<element>format</element> <qualifier>extent</qualifier>
Identificatore della risorsa	Stringa di caratteri o numero univoci di identificazione della risorsa: ad es. l'URL o l'URN	<element>identifier</element> <qualifier>sici</qualifier>
Fonte	Stringa o un numero univoci di identificazione della fonte da cui la risorsa è derivata	<element>source</element> <unqualified>
Lingua	Lingua del contenuto della risorsa nella forma standard ANSI NISO Z39.53	<element>language</element> <unqualified>
Relazione	Relazione della risorsa con altre risorse	<element>relation</element> <qualifier>isformatof</qualifier>

	Copertura	Caratteristiche spaziali e/o temporali della risorsa. Può contenere dati relativi alla localizzazione geografica del soggetto della fotografia in base alla organizzazione amministrativo-territoriale attuale: coordinate, riferimento alla cartografia e quanto è utile per reperire l'esatta collocazione	<code><element>coverage</element></code> <code><qualifier>spatial</qualifier></code>
	Diritti	Link ai dati relativi al copyright o di diritti d'autore della risorsa	<code><element>rights</element></code> <code><qualifier>accessrights</qualifier></code>
Campi aggiuntivi utili alla ricerca storico-geografica ed ecologico-storica LASA	Localizzazione punto di ripresa	Dati relativi al luogo della ripresa del soggetto raffigurato in base alla organizzazione amministrativo-territoriale attuale. Può contenere coordinate, riferimento alla cartografia e quanto è utile per reperire l'esatto punto di ripresa (spot) della fotografia	<code><element>coverage</element></code> <code><qualifier>spot</qualifier></code> <code><scope_note></code> Dati relativi al luogo della ripresa del soggetto raffigurato in base alla organizzazione amministrativo-territoriale attuale. Può contenere coordinate, riferimento alla cartografia e quanto è utile per reperire l'esatto punto di ripresa della fotografia. Luogo preciso dove era posizionata la fotocamera al momento dello scatto <code></scope_note></code> <code></dctype></code> Note: <i>spot</i> non è un qualificatore standard.

Valutazione storica	Questo campo vuole raccogliere le informazioni relative a un uso documentario della fotografia come fonte storica in relazione alla ricerca per la quale è stata selezionata. È un campo aperto che può essere utilizzato per riportare una prima valutazione del documento (ad esempio, per quanto riguarda le dinamiche della vegetazione, si segnalerà la presenza di pratiche agro-silvo-pastorali attive o “indizi” di una loro presenza pregressa, ecc.). Può ovviamente contenere riferimenti o link all’URI (URL o URN) di una o una serie di risorse	<code><element>description</element></code> <code><qualifier>hystorical</qualifier></code> <code><scope_note></code> Questo campo vuole raccogliere le informazioni relative a un utilizzo della fotografia come fonte storica. <code></scope_note></code> <code></dctype></code> Note: <i>hystorical</i> non è un qualificatore standard.
Ripresa di confronto	Collegamento diretto con l’URI (URL o URN) della ripresa di confronto attuale (o viceversa della fotografia storica se si sta schedando una immagine derivata da una ripresa di confronto ottenuta dallo stesso <i>spot</i>)	<code><element>coverage</element></code> <code><qualifier>spot</qualifier></code> Note: <i>spot</i> non è un qualificatore standard.

Tabella 3.2.1: Rappresentazione dello schema di *metadati* mutuato dal Dublin Core adattato allo scopo di accogliere dati utili alla ricerca storico-geografica ed ecologico-storica.

3.3 Le fonti cartografiche

All’interno delle attività del LASA la cartografia, in particolare quella storica, è stata usata da un lato con attenzione al codice usato dai compilatori nella stesura del documento, e dall’altro con attenzione all’incrocio con altre fonti, da quelle osservazionali e “di terreno” a quelle testuali o fotografiche. La natura del documento cartografico ha reso imprescindibili alcune necessità:

- la scelta di uno standard bibliografico di riferimento, per l’aspetto catalogafico;

- la conservazione della terminologia originale della fonte, al fine di preservarne la complessità senza imprecisioni generalizzanti;
- la conservazione del contesto documentario, con la creazione di una serie o una rete di fonti;
- la conservazione, d'altra parte, del collegamento col progetto di ricerca, al fine di rendere quanto più espliciti i criteri di collegamento tra le fonti⁵².

Per quanto concerne il primo punto, una volta assunto come standard di riferimento l'*International Standard Bibliographic Description for Cartographic Material*, si sta al momento lavorando alla creazione di una scheda che la sintetizzi adattandola alla risorsa digitale. Si sta procedendo in modo da modificare DSpace, in virtù della sua versatile natura di Open Source, affinché usi un altro standard di descrizione dell'oggetto digitale.

Il secondo punto, ossia la conservazione della terminologia originale usata dalla fonte, si rende ancor più necessario alla luce delle considerazioni emerse durante i seminari *Storie di cartografi, storie della cartografia: la biografia nella ricerca geografica*⁵³ e *Professione Cartografo*⁵⁴: nella cartografia storica, colui che crea la carta è una figura che ha competenze diverse a seconda del grado, della professione, della mansione svolta o persino della finalità che la carta assolve. Esistono così ufficiali topografi, disegnatori, mappatori, catastatori o agrimensori ma anche, se ci si riferisce a epoche più recenti, operatori di volo: tutte queste figure sono implicate diversamente nella creazione della carta, mantenendone comunque la responsabilità. Questa complessità rischia di perdersi se appiattita sotto una univoca voce "autore": la questione è stata rilevata ma ancora non si è trovata una soluzione che risulti soddisfacente sia per le esigenze di ricerca storica sia per quelle di massima standardizzazione della gestione di risorse digitali.

Alla terza esigenza, cioè alla conservazione del contesto documentario della fonte cartografica, provvedono una serie di collegamenti tra documenti digitali, che permettono di ricostruire la serie di fonti (per esemplificare, minute di campagna prodotte dai topografi di metà Ottocento a scala diversa e a distanza di anni sugli stessi luoghi).

Al quarto e ultimo punto, infine, cioè alla conservazione di collegamenti tra fonti diverse, con criteri che rimangono esterni alla fonte ma stabiliti dal progetto di ricerca, provvedono ugualmente collegamenti e mappature dei documenti: essi possono consistere in rimandi tra fonti cartografiche e fotografiche, o a documenti di altro genere, usati all'interno del progetto come rete di fonti. Tale passaggio permette di rendere percepibile una lettura analitica del documento, resa possibile dal progetto stesso⁵⁵.

3.4 Risultati e problematiche

Il progetto DismecSpace è allo stadio embrionale: molti aspetti sono ancora in fase di progettazione, altri vengono definiti via via che si affinano

gli strumenti e si mettono a fuoco gli obiettivi. Resta fondamentale l'approccio analitico e l'ispirazione alla *topographical history*. Si ripropone quindi, anche a livello dell'organizzazione di un archivio digitale, l'unione relazionale dei dati di terreno con quelli documentario/testuali attraverso l'integrazione, tutta da costruire, tra strumenti diversi come gli *institutional repository* e programmi GIS.

Qualche complicazione sorge però sul lato operativo: la possibilità di instaurare e rendere espliciti collegamenti tra fonti apparentemente lontane conservando il contesto documentario risulta estremamente interessante in fase di ricerca, ma sorgono complicazioni relative alla gestione delle fonti divenute oggetto digitale che impongono di confrontarsi con questioni estremamente dibattute. Tra tutte quella dei diritti di proprietà intellettuale che limita drasticamente la restituzione del lavoro a un'utenza che non sia quella strettamente "interna" andando contro la stessa politica della Unione Europea che sostiene numerosi progetti volti a creare una «biblioteca digitale europea che renda le risorse informative e i contenuti culturali europei più facili e interessanti da utilizzare in un contesto in linea multiculturale e plurilinguistico»⁵⁶.

La seguente bibliografia comprende i principali testi menzionati in nota e altri riferimenti bibliografici significativi:

Auer P., Cavallini F., Giffi E. (a cura di), *Normativa per l'acquisizione digitale delle immagini fotografiche*, Roma 1998.

Carvalho J., *Soluzioni informatiche per microstorici*, in «Quaderni storici», 26 (1991), 78, pp. 761-793.

Bush V., As We May Think, <<http://www.theatlantic.com/doc/194507/bush>> [1945:2007].

Cevasco R., *Nuove risorse per la geografia del turismo rurale, ecologia storica e risorse ambientali nell'appennino ligure-emiliano*, in «Bollettino della Società geografica italiana», 10 (2005), fasc. 2, pp. 345-374.

Chiocchetti F., *Fonti storiche, fonti digitalizzate, metafonti. Qualche nota su alcune iniziative recenti*, in «Storia della storiografia», fasc. 44 (2003), pp. 104-120.

David P. A., *The Consequences for Internet-Mediated Research Collaborations of Broadening IPR Protections*, <<http://emlab.berkeley.edu/users/bhhall/ipconf/David901.pdf>> [2007].

Documento di sintesi del progetto tecnico-scientifico per il Portale della Cultura Italiana del Ministero per i Beni e le Attività Culturali <http://www.otebac.it/siti/realizzare/PICO_2_1sintesi.pdf> [2005:2007].

Dublin Core Metadata Element Set, Version 1.1, <<http://www.dublincore.org/documents/dcmi-type-vocabulary/>> [2007].

Fasella C., *Il modello bibliografico frbr e il Dublin Core*, Seminario internazionale sui metadati, 3 aprile 2001,

- <<http://www.iccu.sbn.it/upload/documenti/Fasella.pdf>> [2001:2007].
- Gill T., Anne J. Gilliland A.J., Woodley M. S., *Introduction to metadata, pathways to digital information*, <http://www.getty.edu/research/conducting_research/standards/intrometadata/index.html> [2006].
- Ginzburg C., *Conversare con Orion*, in «Quaderni storici», 34 (2001), fasc. 108, pp. 905-914.
- Grendi E., *Storia di una storia locale. L'esperienza ligure 1792-1992*, Venezia 1996.
- Itzcovich O., *Dal mainframe al personal computer nella storiografia quantitativa*, in *Storia & computer, alla ricerca del passato con l'informatica*, a cura di S. Soldani, L. Tomassini, Milano 1996.
- Gruppo MAG, Biblioteca Centrale Nazionale di Firenze, *Appunti per la definizione di un set di metadati gestionali-amministrativi e strutturali per le risorse digitali, versione 0 1*, <<http://www.bncf.firenze.sbn.it/progetti/mag/MetaAGVZintroduzione.PDF>> [2002:2006].
- Lo storico e il documento digitalizzato*, in «La Berio», 26 (2005), fasc. 1.
- L'uso del calcolatore in storiografia*, Milano 1993.
- Lerner J., Triole J., *The simple economics of open source*, <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=224008> [2000:2003].
- Leroi-Gourhan A., *Il gesto e la parola*, Torino 1977.
- Pierazzo E. (a cura di), *MAG metadati amministrativi e gestionali, manuale utente, versione: 2.0.1*, <<http://www.iccu.sbn.it/upload/documenti/manuale.html>> [2006:2007].
- Magliano C., *Introduzione al gruppo di lavoro sui metadati descrittivi*, Seminario internazionale sui metadati, 3 aprile 2001, <<http://www.iccu.sbn.it/upload/documenti/Magliano.pdf>> [2001:2006]
- Lo standard nazionale dei metadati gestionali amministrativi*, in «Digitalia», 0 (2005).
- Mackenzie S. (a cura di), *Dspace, an open source dynamic digital repository*, in «d-lib magazine», <<http://www.dlib.org/dlib/january03/smith/01smith.html>> [2003:2006].
- Minervini E., *Prime proposte per il trattamento del materiale fotografico in ambito archivistico*, <<http://www.biblioteche.regione.lombardia.it/regarc/foto.pdf>> [2002:2006].
- Mokyr J., *I doni di Atena. Le origini storiche dell'economia della conoscenza*, Bologna 2004.
- Moreno D., *Ecologia storica*, in *La storiografia italiana degli ultimi vent'anni - età moderna*, a cura di L. De Rosa, Bari 1989.
- Moreno D., *Dal documento al terreno. Storia e archeologia dei sistemi agro-silvo-pastorali*, Bologna 1990.
- Moreno D., Montanari C., *The use of historical photographs as source in the*

study of dynamics of vegetational groups and woodland landscape, in F. Salbitano (a cura di), *Human influence on forest ecosystems development in Europe*, Bologna 1989.

Oder N., *Google scholar links with libs*, in «Library Journal» <<http://libraryjournal.com/article/ca516043.html>> [2005:2007].

Ortoleva P., *Persi nella rete? Circolazione del sapere storico*, in *Storia & Computer, alla ricerca del passato con l'informatica*, a cura di S. Soldani, L. Tomassini, Milano 1996.

Petrella M., Santini C., *Risorse in rete per l'iconografia della città europea in età moderna e contemporanea. Un modello per la valutazione dei siti e dei materiali disponibili*, in «Storia e futuro», <<http://www.storiaefuturo.com>> [1996:2006].

Rowland R., *L'informatica e il mestiere dello storico*, in «Quaderni storici», 26 (1991), 78.

Santoro M., *Dall'analogico al digitale, la conservazione dei supporti non cartacei*, in «Biblioteche oggi», 19 (2001), 2.

Thaller M., *Possiamo permetterci di usare il computer? Possiamo permetterci di non usarlo?*, in «Quaderni storici», 20 (1985), fasc. 60, pp. 871-890.

CLIO, *A Database System* <<http://ftp.univie.ac.at/packages/kleio/pre-v8/doc/hgrb11.tar.Z>> [1993:2007]

Using Dublin Core, <<http://www.dublincore.org/documents/usageguide/>> [2007].

Woollard M., Denley P., *Source-Oriented Data Processing for Historians: a Tutorial for CLIO*, <<http://ftp.univie.ac.at/packages/kleio/pre-v8/doc/hgra23.tar.Z>> [1993:2007].

Vitali S. *Passato digitale*, Milano 2004.

Verkaik R., *Wikipedia and the art of censorship*, in «The Independent», <http://news.independent.co.uk/sci_tech/article2874112.ece> [2007:2007].

Wikipedia, anche Cia e Vaticano "ritoccano" le voci scomode, in «La Repubblica», <http://www.repubblica.it/2006/06/sezioni/scienza_e_tecnologia/wikipedia-2/cia-vaticano/cia-vaticano.html> [2007].

¹ Si ringrazia M. T. Sanguineti (Biblioteca Universitaria di Genova, membro del Gruppo di lavoro per l'elaborazione di linee guida per la digitalizzazione di documenti cartografici - ICCD) che ha fornito suggerimenti bibliografici e collaborazione, alla luce dell'esperienza di digitalizzazione del materiale cartografico in corso presso la Biblioteca Universitaria di Genova. L'articolo è frutto di lavoro e riflessioni comuni.

² A. Leroi-Gourhan, *Il gesto e la parola*, Torino 1977, II, pp. 303-310.

³ Per Bush la macchina avrebbe provocato non solo la nascita di nuove figure professionali, i *trail blazers* (battistrada), specializzate nel creare nuovi percorsi associativi tra la grande massa di informazioni esistenti, ma anche la diffusione di nuovi tipi di enciclopedie caratterizzate da una rete già stabilita di percorsi, che però l'utente avrebbe potuto arricchire a proprio

piacimento: V. Bush, *As We May Think*, <<http://www.theatlantic.com/doc/194507/bush>> [2007].

⁴ È interessante notare che il Memex è stato fonte di ispirazione dell'On-Line System (NLS), un sistema *software* progettato da Douglas Engelbart alla fine degli anni Sessanta, il primo esempio funzionante di *groupware*, ovvero un sistema pensato per fornire un supporto per il lavoro cooperativo e la condivisione di informazioni tra *computer*.

⁵ Per Joel Mokyr, «alla vigilia della rivoluzione industriale, [...] una brama diffusa di conoscenza [...] penetrò in Gran Bretagna fin nelle più piccole cittadine del regno, dove erano molto richiesti i conferenzieri itineranti». Queste conferenze attiravano un pubblico numeroso ed entusiasta che si riuniva per assistere a dimostrazioni sperimentali degli ultimi ritrovati della tecnica. E nella maggior parte dei casi gli inventori erano abili e intraprendenti artigiani assetati di innovazione e che avrebbero fatto prosperare le “microinvenzioni” incrementali, ovvero il graduale perfezionamento delle idee pionieristiche. J. Mokyr, *Leggere la rivoluzione industriale*, Bologna 2002, pp. 63-66.

⁶ Op. cit., p. 119.

⁷ P.A. David, *Communication norms and the collective cognitive performance of “Invisible Colleges”*, in G. Barba Navaretti et al., *Creation and Transfer of Knowledge: Institutions and Incentives*, New York e Berlin 1997, citato in J. Mokyr, *I doni di Atena. Le origini storiche dell'economia della conoscenza*, Bologna 2004, p. 55.

⁸ P.A. David, *The Consequences for Internet-Mediated Research Collaborations of Broadening IPR Protections*. Atti del convegno *Conference on the Law, Economics and History of Intellectual Property*, Berkeley (California), 5-6 ottobre 2001, <<http://emlab.berkeley.edu/users/bhhall/ipconf/ipconfprogram.html>> [2001:2007], p. 5.

⁹ In ambito informatico gli esempi potrebbero essere davvero molti; indichiamo qui soltanto i centri più popolari: SourceForge, <<http://www.sourceforge.net>>; FreshMeat, <<http://www.freshmeat.net>>; Apache Software Foundation, <<http://www.apache.org>>; GNU, <<http://www.gnu.org>>; Free Software Foundation, <<http://www.fsf.org>>; The Code Project <<http://www.codeproject.com/>>; CodeGuru, <<http://www.codeguru.com/>>.

¹⁰ È interessante rilevare che, sebbene molti autori si siano affrettati nel definirlo di volta in volta un modello economico del tutto nuovo, oppure basato su un'improbabile economia del dono (attingendo esempi dal vasto patrimonio etnografico a nostra disposizione), l'Open Source si fonda in ogni caso su di una tradizione di scambio e sviluppo collettivo dei programmi che caratterizza l'informatica sin dagli esordi. L'esempio più antico di comunità organizzata di utenti impegnati nello scambio gratuito di codice sorgente e informazioni, è costituita senza dubbio da SHARE, fondata nel 1952 da IBM. Cfr. M. Campbell-Kelly, *From Airline Reservation to Sonic the Hedgehog: a history of the software industry*, Cambridge (Mass.) 2004, pp. 31-34.

¹¹ Ne è un ottimo esempio WikiScanner, uno strumento nato dalla necessità di mantenere un certo controllo sull'accuratezza delle voci dell'enciclopedia in modo da eliminare, nei limiti del possibile, modifiche arbitrarie. Si veda: R. Verkaik, *Wikipedia and the art of censorship*, in «The Independent», <http://news.independent.co.uk/sci_tech/article2874112.ece> [2007:2007]; *Wikipedia, anche Cia e Vaticano “ritoccano” le voci scomode*, in «La Repubblica», <http://www.repubblica.it/2006/06/sezioni/scienza_e_tecnologia/wikipedia-2/cia-vaticano/cia-vaticano.html>; [2007:2007]. Inoltre su WikiScanner: *WikiScanner*, in «WikiPedia», <<http://en.wikipedia.org/wiki/WikiScanner>> [2007:2007].

¹² L'esempio più illustre è forse rappresentato dalla ricerca condotta sul Catasto fiorentino del 1427 di Herlihy e Klapisch, elaborato tra la fine degli anni Settanta e l'inizio degli anni Ottanta da D. Herlihy e Christiane Klapisch-Zuber; *Florentine Renaissance Resources: Online Catasto of 1427* <<http://www.stg.brown.edu/projects/catasto/overview.html>> [2003:2006].

¹³ R. Rowland, *L'informatica ed il mestiere dello storico*, in «Quaderni storici», 26 (1991), 78, p. 698.

¹⁴ S. Vitali, *Passato digitale*, Milano 2004, p. 39.

¹⁵ M. Woollard e P. Denley, *Source-Oriented Data Processing for Historians: a Tutorial for CLIO*, <<http://ftp.univie.ac.at/packages/kleio/pre-v8/doc/hgra23.tar.Z>> [1993:2007], pp. 45-47.

¹⁶ M. Thaller, *Possiamo permetterci di usare il computer? Possiamo permetterci di non usarlo?*, in «Quaderni storici», 20 (1985), 60, p. 872.

¹⁷ Per esempio, ipotizziamo di possedere un database costituito da nomi, cognomi e città. Se noi

desiderassimo l'elenco di tutti coloro che non sono nati a Roma, nell'interfaccia a caratteri di CLIO, dovremmo digitare quanto segue:

```
query name=censsamp;part=relp:birthto=not "Roma"
part=:surname
stop
```

Dove: "query" indica l'azione da eseguire (in questo caso l'interrogazione del database); "name" rappresenta il nome del gruppo di appartenenza dei dati che stiamo elaborando, mentre "part" indica il gruppo che dipende da quello definito in "name"; "birthto" rappresenta la posizione dentro al gruppo, ovvero il campo del record che contiene il valore. In questo caso "birthto=not Roma" significa letteralmente: "cerca tutti i valori che nella posizione indicata da birthto non riportano Roma". Delle ultime due, la prima istruzione ("part=:surname") indica i dati che desideriamo visualizzare, cioè il cognome, mentre l'ultima di fatto chiude l'operazione. Per un elenco completo dei comandi di CLIO, si veda: M. Thaller, *CLIO, A Database System* <<http://ftp.univie.ac.at/packages/kleio/pre-v8/doc/hgrb11.tar.Z>> [1993:2007].

¹⁸ Per una corretta ed esaustiva definizione di *institutional repository*, A. M. Tammara, *Ruolo e funzionalità dei depositi istituzionali*, <<http://www.bibliotecheoggi.it/2004/20041000701.pdf>> [2004:2007].

¹⁹ Propriamente, i *metadati* costituiscono un insieme di informazioni relative alle caratteristiche dei file all'interno di un *filesystem*, in modo che il sistema operativo sappia come trattarli. Attualmente esistono diversi standard e ognuno è stato pensato a scopi diversi. Il più diffuso, anche nelle diverse varianti, è senza dubbio il Dublin Core, sviluppato dall'OCLC (On line Computer Library Center), la rete dei servizi bibliotecari americana, la cui qualità principale è costituita dalla grande semplicità d'uso. Si veda: Dublin Core Metadata Initiative, <<http://www.dublincore.org/>> [1995:2007].

²⁰ Si veda per esempio CDSWare prodotto dal CERN, <<http://cdsware.cern.ch/invenio/index.html>>; Eprints: <<http://www.eprints.org/software/>>; Gnuteca: <<http://www.gnuteca.org.br/>>; Greenstone: <<http://www.greenstone.org>>; Dspace: <<http://www.dspace.org>>. Un elenco più o meno completo può essere trovato qui: <http://www.unesco.org/cgi-bin/webworld/portal_freesoftware/cgi/page.cgi?d=1&g=Software/Digital_Library/index.shtml>.

²¹ R. Cerri, *L'automazione degli archivi storici*, in *Storia & Computer*, a cura di S. Soldani, L. Tomassini, Milano 1996, p. 126.

²² A. De Robbio, A. Marchitelli, *E-LIS EPrints for Library and Information Science: un modello di archivio per biblioteche digitali aperte*, in *Archivi aperti e comunicazione scientifica*, a cura di A. De Robbio, p. 139 <<http://www.storia.unina.it/cliopress/derobbio.htm>> [2007:2007].

²³ Cerri, *L'automazione degli archivi storici* cit., p. 128.

²⁴ Per una descrizione completa e dettagliata degli elementi Dublin Core, si veda: *Dublin Core Metadata Element Set, Version 1.1* <<http://www.dublincore.org/documents/dces/>> [2007].

²⁵ Non bisogna dimenticare che in linea di massima, l'utilizzo del Dublin Core avviene per mezzo di una rappresentazione in formato XML, maggiormente comprensibile ai *computer*. Secondo questa notazione, i campi descritti potrebbero presentarsi sotto la seguente forma: «<dc:value element="coverage" qualifier="spatial">». Quindi l'utilizzo di una notazione *elemento.qualificatore*, è qui utilizzata per comodità e chiarezza.

²⁶ *Documento di sintesi del progetto tecnico-scientifico per il Portale della Cultura Italiana del Ministero per i Beni e le Attività Culturali* <http://www.otebac.it/siti/realizzare/PICO_2_1sintesi.pdf> [2005:2007].

²⁷ La versione in uso presso il Dipartimento di Storia Moderna e Contemporanea è la 1.3, e a essa si riferisce il presente paragrafo. Per maggiori informazioni su DSpace, si veda <<http://www.dspace.org>>.

²⁸ Per un elenco dei depositi istituzionali operativi, si veda: *Who's Using DSpace?* <<http://wiki.dspace.org/index.php/DspaceInstances>> [2007].

²⁹ Come per esempli la quantità di programmatori attivamente coinvolti nel progetto (quattordici), il numero di *download* e via dicendo. Si veda: *DSpace project*, <<http://sourceforge.net/projects/dspace/>> [2007].

³⁰ Elementi informativi eterogenei, cioè costituiti da documenti diversi, ma strettamente vincolati tra loro e spesso caratterizzati dalla presenza di elementi costituiti da formati diversi

(testo, immagini, filmati). T. Gill, A. J. Gilliland, M. S. Woodley, *Introduction to metadata, pathways to digital information* <http://www.getty.edu/research/conducting_research/standards/intrometadata/index.html> [2007].

³¹ Al momento DSpace è predisposto per indicizzare un numero limitato di elementi Dublin Core: autore, titolo, argomento, *abstract*, serie, identificatore e lingua (secondo la denominazione Dublin Core, rispettivamente: *contributor*, *title*, *subject*, *description.abstract*, *relation.ispartofseries*, *identifier* e *language.iso*).

³² Si veda nota 31.

³³ Conservato nel *database*, del quale DSpace rappresenta lo strumento di gestione, come illustrato nel paragrafo 1.5.

³⁴ J. Carvalho, *Soluzioni informatiche per microstorici*, in «Quaderni storici», 26 (1991), 78, p. 777.

³⁵ Un esempio è costituito da RDF (Resource Description Framework), il quale è: «a foundation for processing metadata; it provides interoperability between applications that exchange machine-understandable information on the Web. RDF emphasizes facilities to enable automated processing of Web resources. RDF can be used in a variety of application areas; for example: in *resource discovery* to provide better search engine capabilities, in *cataloging* for describing the content and content relationships available at a particular Web site, page, or digital library, by *intelligent software agents* to facilitate knowledge sharing and exchange, in *content rating*, in describing *collections of pages* that represent a single logical “document”, for describing *intellectual property rights* of Web pages, and for expressing the *privacy preferences* of a user as well as the *privacy policies* of a Web site. RDF with *digital signatures* will be key to building the “Web of Trust” for electronic commerce, collaboration, and other applications»; *Resource Description Framework (RDF): Model and Syntax Specification*, <<http://www.w3.org/TR/1999/REC-rdf-syntax-19990222/>> [1999:2007].

³⁶ Op. cit., p. 778.

³⁷ *L'Archivio Ligure della Scrittura Popolare*, <http://www.dismec.unige.it/laboratori_centri/ALSP/chisiamo.htm> [2007].

³⁸ O. Itzcovitch, *L'uso del calcolatore in storiografia*, Milano 1993, p. 41.

³⁹ Op. cit., p. 42.

⁴⁰ Come descritto al punto 2 dell'elenco riportato nel paragrafo 1.6, al momento DSpace non possiede un *database* che potremmo definire *di lavoro*, ossia visibile esclusivamente dall'utente e perciò separato da quello centrale, nel quale il ricercatore può effettuare liberamente modifiche e inserimenti senza per questo influire sul patrimonio complessivo di informazioni, se non inseguito a un deliberato inserimento.

⁴¹ Se l'*input* è costituito da un *file XML* opportunamente organizzato, al suo interno il dispositivo è in grado interpretarlo inserendo i dati ricavati nelle giuste tabelle del proprio *database*. Inoltre è possibile il processo inverso, ovvero l'estrazione dei dati e il loro salvataggio nello stesso formato (figura 2.1.9).

⁴² Ovviamente questo procedimento può interessare solo gli elementi i cui valori sono ricorrenti, mentre quelli appartenenti al singolo elemento, vengono di fatto ignorati. Tuttavia nel caso specifico, il processo è stato pensato per effettuare la correzione delle impostazioni iniziali, ossia quelle relative ai valori desunti dalla scheda della carta contenuta nel vecchio sistema.

⁴³ La deperibilità di questi supporti, è nota; la loro conservazione richiede cura e copie periodiche dei contenuti. Per un'analisi dettagliata dei problemi relativi alla conservazione dei dati digitali, si veda O. Itzcovitch, *Lo storico e il documento digitalizzato*, in «La Berio», 26 (2005), 1.

⁴⁴ L'inserimento dei dati in un nuovo *database* richiede in ogni caso procedure adeguate, la cui elaborazione però risulterebbe semplificata proprio dalla natura ordinata e strutturata dei documenti esportati.

⁴⁵ Tale impostazione è debitrice sul piano metodologico, della *topographical history* o *local history* anglosassone (tradotta in italiano come «storia locale» o «topografica»; cfr. E. Grendi, *Storia di una storia locale. L'esperienza ligure 1792-1992*, Venezia 1996). Si tratta di un metodo che può essere accostato allo sviluppo dell'analisi microstorica all'interno della storiografia economica e sociale. Quella che, nella pratica della ricerca, si configura come una «microanalisi storico-geografica della società territoriale e del suo contesto topografico» (cfr. R. Cevasco,

Nuove risorse per la geografia del turismo rurale, ecologia storica e risorse ambientali nell'appennino ligure-emiliano, in «Bollettino della Società geografica italiana», 10, 2005, pp. 345-350) si basa su un «paradigma individualizzante» basato su alcuni principi generali. Tra questi: 1) «l'approccio regressivo» (consente a livello di sito di registrare, al momento dell'osservazione, un preciso stadio, il più recente, delle dinamiche ambientali in atto); 2) «l'osservazione dal particolare al generale» (che consente di introdurre il concetto di "traccia documentaria" – adottando nella ricerca storica il «paradigma indiziario» delineato da Carlo Ginzburg – e permette una più facile messa in evidenza di eventuali "discontinuità" nei processi/dinamiche ambientali; 3) «una forte contestualizzazione delle evidenze» (una particolare attenzione alla produzione e valutazione delle fonti documentarie e osservazionali). Il metodo è adottato con successo soprattutto nell'analisi delle dinamiche della vegetazione condotte col metodo dell'ecologia storica (cfr. D. Moreno, *Ecologia storica*, in *La storiografia italiana degli ultimi vent'anni. Età moderna*, a cura di L. De Rosa, Bari 1989, pp. 402-413). Dallo studio delle dinamiche ambientali generali (per stadi evolutivi di progressione/regressione attorno a uno stato di equilibrio ideale), si raggiunge l'individuazione dei processi ambientali reali, cioè storici, che si sono svolti e sono registrabili nell'ecologia del sito (vegetazione, suolo, popolamenti animali, habitat, ecc.). Presso il Dipartimento di Storia Moderna e Contemporanea dell'Università di Genova è in attività da oltre un decennio un Seminario Permanente di Storia Locale all'interno del quale i problemi metodologici sono ampiamente dibattuti e ai cui lavori molti dei componenti del gruppo di lavoro che fa capo al LASA fanno riferimento.

⁴⁶ Per il triennio 2004-2006 si possono ricordare nell'ordine: 1) il progetto *ZUM: archeologia e storia ambientale per la conoscenza, conservazione e gestione delle zone umide liguri* (Direzione Regionale per i Beni Culturali e Paesaggistici della Liguria; Programma Europeo *Cultura 2000* – azione 2: *Cultural Cooperation Agreements «Our Common European Cultural Landscape Heritage»* (ECL) coordinato dall'Università di Bergen (Norway): *Per un Centro di interpretazione del patrimonio rurale ed ambientale delle valli del Penna*; Doc.U.P. – Obiettivo 2 (2000-2006) decisione della Commissione Europea C (2001) 2044 del 7/9/2001 - Misura 2.6 – *Attività di supporto alla gestione ambientale regionale componente b) realizzazione della Rete Natura 2000*: 2) il progetto «*Boschi e Biodiversità*» (in collaborazione col Parco Naturale Regionale dell'Aveto, in provincia di Genova) e 3) il progetto «*Roccagrande, la storia dell'uomo e della natura - valorizzazione del sito Roccagrande M. Pù*» (su commissione della Comunità montana Val Petronio, in provincia di Genova).

⁴⁷ Le fonti dette "di terreno" (o di campagna, risultato del *fieldwork* dell'etnografo, dell'archeologo, del geobotanico, del geografo-storico), sono da considerarsi come le "evidenze" raccolte con i metodi delle diverse specializzazioni disciplinari. Debbono ovviamente essere valutate ed elaborate quali fonti storiche, cambiando dunque, almeno in parte, il proprio statuto originario.

⁴⁸ Si veda <<http://dismecspace.unige.it>>.

⁴⁹ D. Moreno, *Dal documento al terreno. Storia e archeologia dei sistemi agro-silvo-pastorali*, Bologna 1990, pp. 56-61.

⁵⁰ Ad aver richiamato l'attenzione degli studiosi è infatti la presenza di un deposito biostratigrafico: un bacino lacustre, oggi deposito fangoso completamente disseccato. Il Lago della Nava aveva conservato una portata significativa, seppure in forma ridotta e stagionale, almeno fino alla metà degli anni Novanta del secolo scorso, in conseguenza delle storiche attività di gestione delle acque connesse allo sfruttamento dei pascoli. Primi risultati delle analisi condotte sono stati presentati al 4° *Congres international d'archeologie medievale e moderne - Medieval Europe*, Parigi, Sorbona, 3-8 settembre 2007.

⁵¹ <<http://dismecspace.unige.it/dspace/handle/123456789/63>> [26-07-2007].

⁵² Secondo gli standard di valutazione delle risorse web per le discipline umanistiche, il progetto che sta alla base di una risorsa digitale dovrebbe essere reso esplicito quanto più possibile per garantire la trasparenza dell'informazione: M. Petrella, C. Santini, *Risorse in rete per l'iconografia della città europea in età moderna e contemporanea. Un modello per la valutazione dei siti e dei materiali disponibili*, in «Storia e futuro», <<http://www.storiaefuturo.com/articoli.php?id=25>> [1996:2006].

⁵³ *Storie di cartografi, storie della cartografia: la biografia nella ricerca geografica*, 8-9 giugno 2006, Torino.

⁵⁴ *Professione Cartografo*, 22 giugno 2006, Genova.

⁵⁵ Si richiama a questo proposito *The Spectator Project*, preso in esame da Chiocchetti in una lettura critica delle fonti storiche e fonti digitalizzate. Nel caso in questione, un progetto di digitalizzazione del periodico settecentesco fa ricorso a collegamenti e ipertesto per evidenziare il ruolo di modello svolto da «The Spectator» nei confronti di altri periodici coevi, effettuando un confronto tra brani. Nell'ambito dell'articolo citato, Chiocchetti riprende da Genet il concetto di *metafonte*, intesa come mutamento della natura epistemologica della fonte in seguito al trattamento informatico. Si tralascia in questa sede tale definizione, accettando tuttavia il concetto per cui «la scelta di un medium non tradizionale» rappresenta «un valore aggiunto per l'iniziativa intrapresa»: F. Chiocchetti, *Fonti storiche, fonti digitalizzate, metafonti. Qualche nota su alcune iniziative recenti*, in «Storia della storiografia», 44 (2003), pp. 104-120.

⁵⁶ 2689^a sessione del Consiglio Istruzione, gioventù e cultura, Bruxelles, 14-15 novembre 2005, <<http://europa.eu/rapid/pressReleasesAction.do?reference=PRES/05/284&language=it>>